

Springer Handbooks in Languages
and Linguistics

SPRINGER NATURE
Reference



Alireza Korangy
Karim Bensoukas
Editors

The Handbook of Berber Linguistics



Springer

Springer Handbooks in Languages and Linguistics

Series Editor

Alireza Korangy, American University of Beirut, Beirut, Lebanon

Springer Handbooks in Languages and Linguistics is the largest and most comprehensive undertaking of its kind in the history of linguistics and the long history of linguistics scholarship. This series of reference works offers a comprehensive survey of all the world's languages—and dialects therein—edited by leading scholars in their respective fields. The handbooks not only reframe prior scholarship in accordance with new findings, but also offer novel insights into each language branch and language family with a mindful observation of both new sources and historical precedence. In encompassing the broad spectrum of contemporary and historical linguistics, and engaging with the structure and function of different languages, contributors excavate a wide range of intellectual linguistic issues from the Middle East, to North and South Africa, Asia, Europe, North America and Oceania. A magnum opus within linguistics, presenting global coverage and boasting extensive geographic diversity from within the corresponding research communities, this collection is the ultimate companion for linguists and language enthusiasts, as well as anthropologists, philologists, poeticians, rhetoricians, and literati, everywhere.

Alireza Korangy • Karim Bensoukas
Editors

The Handbook of Berber Linguistics

With 69 Figures and 194 Tables

 Springer

Editors

Alireza Korangy
American University of Beirut
Beirut, Lebanon

Karim Bensoukas
Mohammed V University in Rabat
Rabat, Morocco

ISSN 2731-068X

ISSN 2731-0698 (electronic)

Springer Handbooks in Languages and Linguistics

ISBN 978-981-99-5689-0

ISBN 978-981-99-5690-6 (eBook)

<https://doi.org/10.1007/978-981-99-5690-6>

© Springer Nature Singapore Pte Ltd. 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.

The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

If disposing of this product, please recycle the paper.

*Alireza Korangy dedicates this book to Iran
Ghazal Korangy, his daughter: “my ad
infinitum, my all, my immortal beloved”
Karim Bensoukas dedicates this book to
Malika, Adam, and Hanae. And to “the
memory of my brother Mohamed”*

Introduction

Studies in Berber linguistics have much room to flourish in terms of interest in the field. Many, from student of linguistics to fully accomplished linguists, are yet to understand and be aware of the wonder that is Berber. Certainly, many scholars whose brilliant work has adorned the pages of this volume are on the other end of that spectrum and deserve the lion's share of credit for maintaining an ongoing and rigorous discourse in Berber languages and linguistics. That said, this work tries to help, in its own way, to bring to purview this fascinating field of study with the help of the very same aforesaid scholars and some younger ones who will be the future flagship of this field of study. The reasons why Berber linguistics has experienced a less-to-be-desired fortune in terms of interest rates a further discussion but certainly politics, immigration laws and issues, and a peculiar lack of advertisement for it are not small reasons. What I can say is that the initiation of this project has much to do with my dear departed mentor and Ph.D. advisor Wolfhart P. Heinrichs of Harvard University whose love for Semitic and Afroasiatic languages and linguistics knew no bound. In fact, his Mouton Grammar Library collection, among his many grammar books, and the thousands of marginal notes and comments he had made in these works would even say that he simply loved language for it is, well, expression, analysis, and comprehension: without language, proper language, thought is nothing but a passing fancy lost in the conundrum of other unrelenting thoughts. That said, his first love was semitic and Afro-Asiatic and through many discussions with dear Wolfhart it became clear to me long ago that there was a need for a large work that could serve as a general platform for Berber, although there were and are many singularly brilliant works in the field – and even more brilliant Berberists – which deal with the macro of Berber linguistics: a thematic and detailed large work became my goal. One would be remiss not to mention the dozens of works meticulously edited and authored by wonderful scholars at Rüdiger Köppe Verlag's Berber Studies Series. I am not sure we have achieved our larger-than-life dream here fully, but we know we have started it with gusto here and we hope it will lead to future works of the same size with even more novel investigations. What we opined for this volume were both micro and macro engagements that can somehow bring together the significance of Berber and encourage further studies. We have done that.

The range of focus of the chapters within the volume is quite extensive, as this volume was always meant to serve as the absolute largest possible platform to date for an extensive scrutiny of Berber. The general themes covered in this volume are Phonology and Phonetics, Morphology and Syntax, Semantics and Pragmatics, Sociolinguistics and Dialectology, Language Teaching and Psycholinguistics, Lexicology, Language Contact and comparative linguistics, Historical Linguistics and Etymology. Under the auspices of the above motley of topics, we have ascertained a secondary necessity in order to better address the more general topics – not to mention paucity of further studies on such sub-themes as Prosody (Stress, Intonation...), Ideophones (and Expressive Language in General), Morpho-Syntactic Categories, Studies in Proto-Berber, Berber and Semitic, Sociolinguistic variation, and several seminal interdisciplinary explorations.

Representativeness of the variety of Berber dialects, Berberist scholars, and linguistics trends that has been adopted is structurally coherent to our end. The chapters of the volume reflect the diversity of Berber varieties and an up-to-date range of scholarship by Berberists – and the resulting linguistic trends. First, as far as the varieties are concerned, the papers deal with the Figuig variety, Kabyle, Senhaja, Siwa, Standard Moroccan Amazigh (Berber), Tamazight, Tarifit, Tashlhit, Touareg, Tunisian Berber, Znaga, as well as Proto-Berber. Concomitantly, a large geographical span is covered, including Algeria, Egypt, Libya, Mali, Mauritania, Morocco, and Tunisia. In addition, the contributors are diverse, and as such are the schools of linguistics in which they work. The volume includes studies by scholars from America, Africa, Australia, and Europe. The editors have also particularly solicited scholars from the dominantly Berber-speaking countries – Algeria and Morocco – and their diaspora. Second, the chapters of the volume assume different schools of linguistics and different theories, ranging over European Structuralism and Generativism and more specific theories like the CV model of phonology and morphology and Optimality Theory to name a few. The editors always felt that without soliciting both English and French contributions, a sizeable vacuum would remain as some of the very top scholars in the field are French speaking. We also feel the presence of the French language as a scholarly medium would add another important dimension in terms of the sources that will be presented in the volume. In fact, in a future work on Afro-linguistics, the same attitude has been maintained.

The Handbook of Berber Linguistics, as mentioned, has a vast scope that is best highlighted through the synopses of its chapters. This is where I would give a snapshot of each chapter; however, I won't as our authors have each given such a snapshot at the beginning of each chapter making such a task redundant here, particularly if it means it will delay your reading of the works herein. As will be seen in the chapters, the volume encompasses a large swath of geography and consequently linguistic and thematic area in its investigations: studies that range from phonetics, toponymy, typology, and onomastics to comparative investigations pertaining to the Japanese language not only expand upon the field as per a linguistic phenomenon but also a cross-linguistic one. We hope you will enjoy reading this

work as we have enjoying pouring over every single word in it with pleasure and utter excitement. As a forever-student of language, it was indeed a huge pleasure learning from every single author in this volume and from my wonderful co-editor and brilliant mensch Karim Bensoukas.

Alireza Korangy

Acknowledgments

I write this acknowledgment as a co-editor on this volume and the editor-in-chief of the series in which it is placed. This volume is a result of sheer rigor on the part of all the contributors who patiently and cordially weathered the many obstacles we faced in finalizing this project. Undoubtedly the first and the biggest thanks go to them who share the spotlight with the unbelievable team at Springer. Alexandra Campbell, you are the backbone of this work, the series, and every success it has had or will ever have. I only feel blessed to have worked with you in this long journey. Swetha Varadharajan's herculean efforts are not to be overlooked as she has made this volume what it is in a palpable sense. Moshika Gaur, thank you for all your support throughout all of the projects in the series and this volume. The Springer team is a dream team for any author indeed. Many thanks go to the other editor of this volume Dr. Karim Bensoukas, whose expertise in the field is the guiding light of this project in all of its stages. Without his co-editorship, this book would only be a dream and would have never realized being a book. We would like to thank our esteemed reviewers for their rigorous attention to detail. This project would have not been what it is if not for them. Last but not least, this work owes its rigor (my side of it) to my daughter Iran Ghazal Korangy: you are my light, my every breath, and the only reason I want to live with hope, produce, and be good at whatever it is I do. You are my purpose, and you are my divine north star: عاشقتم عمرم.

Contents

Section I Berber Phonetics/Phonology	1
1 Durational Correlates of Geminate Stops: The Case of Berber and Japanese	3
Fayssal Bouarourou, Tomoki Koya, Saïd Bouzidi, Béatrice Vaxelaire, and Rudolph Sock	
2 “Within-Category” Consonant–Vowel Interactions in Tashlhit: Labialized Consonant Alternations	25
Karim Bensoukas	
3 Acquisition in a Multilingual Context: English Mid and Low Vowel Contrasts by Native Speakers of Tashlhit	55
Nancy C. Kula and Nabila Louriz	
4 Glide: High Vowel Alternation at the Syntax–Phonology Interface	87
Mohamed Lahrouchi and Rachid Ridouane	
5 The Phonetics and Phonology of Tashlhiyt Geminates: An Overview	103
Rachid Ridouane	
Section II Cultural Linguistics, Sociolinguistics, Psycholinguistics, and Pedagogy	121
6 Le corse et l’amazighe, de la minoration à l’officialisation et de la famille à l’école	123
Bouchra El Barkani	
7 Swadesh in the Sahara: Lexicostatistical Lists in Berber/Amazigh Languages and Their Loan Components	153
Anthony Grant	

8	Auditory Word Recognition in Tashlhit/Berber: Supraliminal Speech Priming	165
	Fatima El Hamdi	
9	Shaping Amazigh Identity: The Case of the Netherlands	195
	Jan Jaap de Ruiter	
10	Language Attitudes Toward Berber: Two Decades After the Official Recognition	211
	Mohammed Errihani	
11	Word Boundaries in the Writing System of Standard Amazigh: Challenges from Tarifit Facebook Users	229
	Naima Tahiri	
12	The Berbers in Tunisia: Mistaken Identity Amid Contentious Politics in Post-Arab Spring Tunisia	255
	Zouhir Gabsi	
	Section III Historical Linguistics and Typology	271
13	L'archaïsme linguistique en berbère (kabyle): étude sur un corpus parémiologique	273
	Abdelaziz Berkai	
14	Reconstructing the Proto-Berber Verb System in Afro-Asiatic Perspective	291
	Vit Bubenik	
15	Proto-Berber Heavy Verbs	307
	Marijn van Putten	
16	Berber <i>isnin</i> and “Defined Numerals”	335
	Vermondo Brugnatelli	
	Section IV Morphology and Syntax (and Typology)	359
17	On Berber Verbs of State	361
	Sabrina Bendjaballah and Martin Haiden	
18	The Status of Berber (Past and Present) and Its Contact with Arabic	385
	Moha Ennaji	
19	Comment la tradition berbérissante a intégré la partie ergative de la syntaxe berbère actuelle dans celle qui est accusative	403
	Abdelaziz Allati	
20	Conditionals in Berber	425
	Ahmed Ech-Charfi	

21	Frequency Norms in Tashlhiyt: A Pilot Study	443
	John Alderete, Jane Li, and Rachid Ridouane	
22	On State and Case in Berber: A Typological Perspective	455
	Aicha Belkadi	
23	Tashlhiyt Berber DP Assembled	493
	Hassan Makhad	
24	Two Waves of Berber Influence on Moroccan Arabic	517
	Jeffrey Heath	
25	Conjugated Adjectives and Participles in Senhaja Berber (Northwestern Morocco)	523
	Evgeniya Gutova	
26	The Use of the Bare Aorist in Figuig Berber (Eastern Morocco)	547
	Maarten Kossmann	
27	Numeral Reduplication in Taqbaylit	571
	Nataša Knežević and Amazigh Bedar	
28	Diatopic Variation in Kabyle Berber	597
	Samir Ben Si Said	
29	The Functions of Independent Personal Pronouns in Berber	615
	Valentina Schiattarella	
30	Action Nouns in Zenaga Berber of Mauritania	637
	Catherine Taine-Cheikh	
Section V	Lexicology and Onomastics	659
31	A Preliminary Account of Hypocoristics in Moroccan Amazigh	661
	Abdelaziz Boudlal and Mohamed Yeou	
32	Comparaison et reconstruction en berbère	683
	Abdallah El Mountassir	
33	Le vocabulaire berbère des plantes. Profondeur historique, conservation et permanence	699
	Mohand Tilmatine	
Index		719

About the Editors



Alireza Korangy received his PhD from the Department of Near Eastern Languages and Civilizations at Harvard University. His field of research is classical Persian and Arabic philology with a special emphasis on poetics, rhetoric, folklore, and linguistics. He has done extensive research and published on Iranian and Persian linguistics. He has also published on Iranian folkloric traditions. Dr. Korangy currently teaches in the Faculty of Humanities and Civilization Studies at the American University in Beirut. He has previously taught at the University of Virginia, the University of Colorado, and Harvard University.



Karim Bensoukas is currently a Senior Lecturer (Professeur de l'Enseignement Supérieur) in the Department of English at the Faculty of Letters and Human Sciences, Mohammed V University in Rabat, Morocco. He has published on linguistic morphology, phonology, the phonology-morphology-syntax interfaces, as well as language contact, variation, and change. His research interests extend to applied sociolinguistics and university pedagogy, particularly the integration of ICT in teaching. He is the editor of *Langues et Littératures*, co-editor of the *International Journal of Arabic Linguistics*, and a member of the advisory board of *Asinag*, *Journal of Applied Language and Culture Studies*, and *Revue des Etudes Amazighes*. He is also “expert-évaluateur” with Centre National pour la Recherche Scientifique et Technique (CNRST), Rabat. Karim Bensoukas is the winner of the 2014 IRCAM-Rabat Award for Research and Scientific Thinking.

Contributors

John Alderete Simon Fraser University, Burnaby, BC, Canada

Abdelaziz Allati Université de Tanger-Tétouan, Tanger-Tétouan, Morocco

Amazigh Bedar Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, Marseille, France

Aicha Belkadi SOAS University of London, London, UK

Samir Ben Si Said Laboratoire Ligérien de Linguistique, Université de Tours, Tours, France

Sabrina Bendjaballah LLING-UMR6310, CNRS & Nantes Université, Nantes, France

Karim Bensoukas Mohammed V University in Rabat, Rabat, Morocco

Abdelaziz Berkai Faculté des lettres et des langues, Département de langue et culture amazighes, Université de Béjaia, Béjaia, Algeria

Fayssal Bouarourou Université de Strasbourg, Strasbourg, France

Abdelaziz Boudlal Faculty of Letters and Humanities Ben M'sik, Hassan II University of Casablanca, Casablanca, Morocco

Saïd Bouzidi Université de Strasbourg, Strasbourg, France
Université Sorbonne Nouvelle, Paris, France

Vermondo Brugnattelli Università di Milano-Bicocca, Milan, Italy

Vit Bubenik Faculty of Arts and Social Science, Memorial University of Newfoundland, St. John's, NL, Canada

Jan Jaap de Ruiter Tilburg University, Tilburg, The Netherlands

Ahmed Ech-Charfi Mohammed V University, Rabat, Morocco

Bouchra El Barkani Institut Royal de la Culture Amazighe, Rabat, Morocco

Fatima El Hamdi Department of English, Mohammed V University in Rabat, Rabat, Morocco

Abdallah El Mountassir Université Ibn Zohr, Agadir, Morocco

Moha Ennaji Sidi Mohamed Ben Abdellah at Fez, International Institute for Languages and Cultures, Fez, Morocco

Mohammed Errihani Purdue University Northwest, Hammond, IN, USA

Zouhir Gabsi Humanities and Social Sciences, Deakin University, Geelong, VIC, Australia

Anthony Grant Edge Hill University, Ormskirk, UK

Evgeniya Gutova University of Navarra (ICS), Pamplona, Spain

Martin Haiden LLING-UMR6310, Nantes Université & CNRS, Nantes, France

Jeffrey Heath University of Michigan, Ann Arbor, MI, USA

Nataša Knežević Nantes Université, Nantes, France

Maarten Kossmann LUCL, Leiden University, Leiden, The Netherlands

Tomoki Koya Université de Strasbourg, Strasbourg, France

Nancy C. Kula Leiden University Centre for Linguistics, Leiden, The Netherlands
University of the Free State, Bloemfontein, South Africa

Mohamed Lahrouchi CNRS & University Paris 8, Paris, France

Jane Li Johns Hopkins University, Baltimore, MD, USA

Nabila Louriz University of Hassan II of Casablanca, Casablanca, Morocco

Hassan Makhad The Polydisciplinary Faculty of Safi, Cadi Ayyad University, Safi, Morocco

Rachid Ridouane Laboratoire de Phonétique et Phonologie, CNRS & Sorbonne Nouvelle, Paris, France

Valentina Schiattarella University of Hamburg, Hamburg, Germany

Rudolph Sock Université de Strasbourg, Strasbourg, France
University of Košice, Košice, Slovakia

Naima Tahiri Université Sidi Mohamed Ben Abdellah, Fés, Morocco

Catherine Taine-Cheikh LACITO (CNRS, Université Sorbonne Nouvelle et Inalco), Paris, France

Mohand Tilmatine University of Cádiz, Cádiz, Spain

Marijn van Putten Leiden University, Leiden, The Netherlands

Béatrice Vaxelaire Université de Strasbourg, Strasbourg, France

Mohamed Yeou Faculty of Letters and Humanities, Chouaib Doukkali University, El Jadida, Morocco

Section I

Berber Phonetics/Phonology

Durational Correlates of Geminate Stops: The Case of Berber and Japanese

1

Fayssal Bouarourou, Tomoki Koya, Saïd Bouzidi, Béatrice Vaxelaire,
and Rudolph Sock

Contents

Introduction	4
Hypotheses	7
Method	7
Results	11
Intersegmental Results for Voiceless Stops	11
Intersegmental Results for Voiced Stops	12
Intrasegmental Results for Voiceless Stops	14
Intrasegmental Results for Voiced Stops	14
Relative Values for Singletons Versus Geminates	15
Discussion and Conclusions	19
Absolute Durational Data	19
Relative Durational Data	20
Concluding Remarks	21
References	23

F. Bouarourou (✉) · T. Koya · B. Vaxelaire
Université de Strasbourg, Strasbourg, France
e-mail: fayssalbouarourou@gmail.com; koyacchi615@hotmail.com; vaxelaire@unistra.fr

S. Bouzidi
Université de Strasbourg, Strasbourg, France
Université Sorbonne Nouvelle, Paris, France
e-mail: saidbzd9@gmail.com

R. Sock
Université de Strasbourg, Strasbourg, France
University of Košice, Košice, Slovakia
e-mail: sock@unistra.fr

Abstract

Berber and Japanese languages are referred to as quantity languages. It is usually posited that closure duration is cross-linguistically the most reliable and consistent acoustic feature of gemination. In this study, we conduct an acoustic analysis of singleton versus geminate voiced and voiceless stops, produced in fast and normal speaking rates, in two varieties of Berber (Tarifit and Taqbaylit) and Japanese. Speech rate is varied in order to evaluate the robustness of the phonological contrast. The main thrust of the present investigation is to show how the singleton versus geminate contrast functions in two related languages, Tarifit Berber and Taqbaylit Berber, which belong to the same Afro-Asiatic family, are compared to Japanese, which belongs to a different language family, the Altaic family. The corpus of elicited materials consists of three different corpora (one corpus for Tarifit, one for Taqbaylit, and one for Japanese) of singleton versus geminate minimal voiced and voiceless stop pairs, inserted in carrier sentences. Acoustic analyses are based on measurements of absolute durations of intersegmental and intrasegmental parameters for the three languages – in the two speaking rate conditions. Relative timing is also observed by examining the proportion of the consonantal duration of the target singleton and geminate consonants within the CV syllable. Results show, at the intersegmental level, that closure duration is the main cue for this phonological contrast in the two speaking rate conditions, cross-linguistically. Intrasegmental results reveal that the acoustic silent phase for voiceless stops and consonantal occlusion for voiced stop are the two acoustic correlates that distinguish the two phonological categories across the three languages. As concerns the other parameters (VTT or Voice Termination Time, VOT or Voice Onset Time and the flanking vowels), they did not show significant differences, except for Japanese where vowels are longer before geminate consonants than before their singleton counterparts.

Keywords

Gemination · Tarifit Berber · Taqbaylit Berber · Japanese · Fast speech · Relative timing

Introduction

Several studies have sought to determine acoustic cues for gemination in many languages (see, e.g., Bouarourou 2014; Bouarourou et al. 2018a, b) on Tarifit Berber. One consistent acoustic characteristic shared by geminates is that their closure is significantly longer than that of their singleton counterpart (see Ridouane 2010 for a review of 24 languages comparing singletons with geminates). Lahiri and Hankamer (1988), for example, investigated the timing properties of singleton/geminate voiceless stops in Turkish and Bengali and showed that closure duration is the most important correlate of the geminate/singleton contrast for both languages. In

addition, VOT is longer for geminates in Turkish, while vowel duration is unaffected. In Bengali, vowel duration is shorter before geminates, but VOT is unaffected. In his acoustic study on gemination in four unrelated languages – Levantine Arabic, Standard Hungarian, Indonesian Madurese, and Swiss-German Bernese – Ham (1998) found that the only acoustic correlate that significantly distinguishes geminates from singletons is closure duration. Positive VOT or burst duration does not contribute to the contrast between these consonants in any of these languages. Other findings suggest, however, that rather than being restricted to durational differences alone, the implementation of gemination may have implications on the spatial organization of the vocal tract, with accompanying spectral consequences. In Malayalam, for example, forms containing geminates differ systematically from those without geminates in terms of phonation, tense versus lax articulations, consonant and vocalic resonances as well as patterns of articulatory variability in adjacent consonants (Local and Simpson 1999). In Japanese, as mentioned in Kawahara (2015) on the phonetics of obstruent geminates, all results in all of the studies cited showed that the duration of geminate consonants is significantly longer than that of their singleton counterparts. He also observed that vowels (V1) are significantly longer before geminates than before singletons. Kawahara (2006) also mentioned that voiced geminates are prohibited in the native vocabulary of Japanese. According to Itô and Armin (1995), voiced geminates are, however, allowed in recent loanwords borrowed from foreign languages (mainly English). They also added that devoicing of these voiced geminates may occur in some instances. Smith (2006) explained that consonants spelled with two letters are often borrowed as geminates (e.g., *slugger* is borrowed as [suraggaa]). This gemination process has created voiced geminates in the loanword phonology of Japanese. A voiced/voiceless distinction is thus contrastive in Japanese geminates. However, Haraguchi (2006) has pointed out that, when voiced obstruent geminates appear with another voiced obstruent, they can undergo optional devoicing (e.g., /guddo/ → [gutto] “good”). In a recent study about Tokyo Japanese, Hussain and Shinohara (2019) found that the devoicing of the word-medial stops was not only observed in voiced geminates, but also that voiced singletons showed devoicing. The duration of the preceding vowel (V1) contributed to distinguishing the voicing contrast in both singleton and geminate stops. The first four spectral moments of C2 stop release bursts did not distinguish the length and voicing contrasts in stops. These results indicate that, although word-medial voiced geminate stops are fully or partially devoiced, the Tokyo Japanese speakers lengthen the preceding vowels (V1) to maintain a voicing contrast. Production patterns of the voiced geminates are considered in relation to marginal or intermediate phonological contrast.

Relative duration of vowel and consonantal units may be distinctive in certain languages since some linguistic contrasts are essentially based on differences in articulatory and/or acoustic durations. Such languages are referred to as quantity languages (see Lahiri and Hankamer 1988) for a leading review of the issue). A vowel quantity language is a language in which short vowels contrast with long vowels; whereas in a consonantal quantity language, single consonants contrast with long or geminate consonants.

One of the advantages in studying such languages lies in the possibilities they offer in analyzing the temporal characteristics of these linguistic contrasts within a common framework of general knowledge on constraints underlying the mechanisms of their production, and their perception. From our point of view, it is of foremost importance to be capable of specifying, in a general manner, the different types of articulatori-acoustic constraints related to the timing of the differences of syllabic structures on which quantity contrasts rely, regardless of the language. We shall, thereafter, try to define means of identifying other factors that may be much more related to linguistic aspects than to idiosyncratic properties of the speech production system (e.g., Sock 2010; Sock and Vaxelaire 2011). Hence, in the first case, one would be more preoccupied with generalizable biomechanical and acoustical constraints, i.e., the physical substrate that constitutes syllable formations for quantity contrasts, cross-linguistically. Whereas in the second case, one would be dealing with language-specific characteristics which, by definition, cannot be captured by any universal linguistic principle.

Varying speech rate (normal and fast speaking rates) experimentally is a useful paradigm for two major reasons. Firstly, it serves as a natural mean of evaluating the robustness of the temporal phonological contrast, i.e., gemination, under investigation here. Indeed, due to the elasticity of speech signals (Gaitenby 1965), the impact of increased speaking rate on the duration of the different acoustic phases, which constitute singleton and geminate syllables, allows appreciating to what extent these syllables and words can be compressed, in terms of ultimate intelligibility. Secondly, increased speaking rate may provoke articulatory gestural reorganizations with remarkable acoustic consequences. The trajectories articulatory-acoustic patterns could be classified, either as patterns pertaining to biomechanical factors inherent to the speech production system, or as motor behavior linked to linguistic viability constraints (Sock 1999; Lehiste 1970).

It is this phenomenon of gestural reorganization, as highlighted by the different acoustic phasing patterns, that is of interest here in this gemination and linguistic viability paradigm. Indeed, we shall try to determine, from our knowledge of linguistic constraints, and within the viability paradigm, how phonetic and phonological contrasts are preserved, despite noticeable spatiotemporal modifications, in a prosodic condition like fast speech rate, usually considered as relatively “difficult.” In the area of human nonspeech, and animal biological behavior, gestural modification by structural simplification obeys basic economy principles, apparently generally operational in movement control. Differentiated motor behaviors are thus neutralized in favor of the most stereotyped ones (see Jeannerod 1988). This is the case in biological behaviors such as pointing, grasping, typing, etc. In a cognitive motor behavior such as speech, simplification processes are, however, very rare, and would only emerge as long as they do not violate linguistic constraints that require perceptual clarity and contrasting categories. Hence, it has been posited by Sock and Vaxelaire (2009) that speech is a special motor behavior.

The main thrust of the present investigation is to show how the singleton versus geminate contrast functions in two related languages: Tarifit Berber and Taqbaylit Berber, which belong to the same Afro-Asiatic family, and Japanese, which belongs

to a different language family, the Altaic family (see for a very detailed literature review on the phonological system of Berber Bouarourou 2014; Bouarourou et al. 2018b; and for Japanese Jeannerod 1988; Kawahara 2006; Kawahara 2015; Koya 2015).

These two unrelated languages were chosen among the different quantity languages that we have previously investigated, because they do not only offer a rationale for studying generalizable and specific linguistic factors (see below), but because they are also part of a programmatic-systematic research that is being carried out on quantity languages in our laboratory.

Hypotheses

It is *hypothesized* that: (1) geminates would have longer closure durations than singletons, as reported in the literature, and especially because of longer articulatory contact extents (lip-to-lip, tongue-to-palate) VTT found for geminates (see Bouarourou 2014; Bouarourou et al. 2010, 2018b). (2) The duration of flanking vowels may be affected by that of geminate consonants: they would be shorter in this environment in case of syllable isochrony. This phonological contrast as a gauge of potentially reinforcing the main cue of consonantal closure should be underlaid by two intrasegmental consonant intervals: (3) acoustic silence for voiceless stops and (4) consonantal occlusion for voiced stops. (5) VOT could be longer for geminates, as their occlusion phase is usually remarkably long, thus retarding onset of voicing, due to high intra-oral pressure. (6) It is likely that VTT (Voice Termination Times) would be longer for singletons than for geminates. Indeed, the shorter the consonantal closure, the higher the proportion of voicing decay within this consonantal closure would take. (7) The speech signal is intrinsically elastic; all segments are expected to undergo compression with increasing speaking rate. This compression of the speech signal should, however, not hinder preservation of phonological contrasts. (8) Emerging acoustic timing patterns, i.e., observed consonantal measures, should be similar for Tarifit and Taqbaylit Berber, two quite related dialects – and different for Japanese, an unrelated language.

Method

Speakers and Corpora

The corpus of elicited materials consists of three different corpora (one corpus for Tarifit, one for Taqbaylit, and one for Japanese) of singleton/geminate minimal pairs, of voiced and voiceless stops, inserted in carrier sentences.

Tarifit: ini ____ i3 umar “say ____ once”

Taqbaylit: inas ____ θikkəlt “say him/her ____ once”

Japanese: are wa ____ san desu “that is Mrs. or Mr. ____ there” (see word list in Table 1.1 below)

Table 1.1 Japanese, Tarifit and Taqbaylit Berber word list

Tarifit		Taqbaylit		Japanese	
Singl.	Gem.	Singl.	Gem.	Singl.	Gem.
amatar	amattar	atas	attas	kapa	kappa
akas	akkas	amakas	amakkas	kata	katta
aqəθ	aqqaθ			kaka	kakka
swadaj	swaddaj	adal	addal	kaba	kabba
itragab	itraggab	amagas	amaggas	kada	kadda
				kaga	kagga

The carrier sentences served to control, as much as possible, the segmental and prosodic context, the number of syllables – and they also give some sense to the uttered sentences.

The Tarifit corpus consists of voiced and voiceless plosives for alveolars /t, d/, for velars /k, g/, and for the uvular /q/ in intervocalic position: they are repeated ten times by six native speakers (two women and four men) from Northern Morocco, at normal and fast speaking rates (the age of the speakers was between 29 and 34 at the time of recording). They are all native speakers of Tarifit Berber with Moroccan Arabic being their second language, and French their third.

The Taqbaylit corpus consists of voiced and voiceless plosives for alveolars /t, d/, and for velars /k, g/ in intervocalic position, and repeated ten times by two native speakers (2 men) from Northern Algeria, at normal and fast speaking rates (the age of the speakers was 36 and 40 at the time of recording). They are all native speakers of Taqbaylit Berber, Algerian Arabic being their second language, and French their third.

The Japanese corpus consists of voiced and voiceless plosives for bilabials /p, b/, alveolars /t, d/, and for velars /k, g/, in intervocalic position, and repeated ten times by four native speakers (two men and two women) of “standard” Japanese, at normal and fast speaking rates (the age of the speakers was 30–35 at the time of the recordings). All the speakers were native speakers of Japanese, with French and English being their foreign languages.

The average age for all speakers in this study was 35 years (± 5 years). Hence, age differences between speakers were kept as minimal as possible. As we are dealing here with a phonological feature, which relies mainly on durational correlates (and much less on spectral differences), gender differences are not considered in this study. Moreover, the speakers of Japanese, Tarifit, and Taqbaylit studied here indeed produce samples that should not essentially differ from random samples, as regards gemination at least, since this phonological contrast entails precise timing constraints in order to be perceptually relevant.

Recordings and Acoustic Measurements

All data were acquired in an anechoic room at the Phonetics Institute of Strasbourg, using highly performant digital equipment and rigorous experimental protocol. The

corpus was recorded with digital recorder Marantz Professional (model PMD 661) and Sennheiser e845S© microphone. The subjects were sitting comfortably 20 cm from the microphone and pronounced sentences ten times at normal and fast speech rates.

Temporal events were detected on the *audio* signal, and specific intersegmental and intrasegmental timing relations between these events allowed determining acoustic durations (ms), which correspond, respectively, to articulatory opening and closing gestures, and to timing between supraglottal and glottal gestures. Thus, for intersegmental timing relations, vowel durations were specified as intervals between onset and offset of a clear formant structure for V1 and V2. Corollary, consonantal closure duration was measured between vowels from offset to onset of clear vocalic formant structures (see Figs. 1.1 and 1.2). As concerns intrasegmental timing relations, VTT (Voice Termination Time) was measured from vowel offset to the last voicing pulse, following voicing decay, within the voiceless plosive (Agnello 1975). Plosive occlusion (i.e., closure duration excluding VOT for voiced plosives), the acoustic silent phase (for voiceless plosives), and VOT (the interval between the burst-release of the plosive and onset of a clear formant structure of the subsequent vowel), were also acquired. It should be noted that following Klatt (1975), VOT could be measured in positive values for both voiced and voiceless plosives (see Figs. 1.1 and 1.2). The choice of VTT as a measure is motivated by the fact that the extent of voicing decay within voiceless consonants could contribute in determining

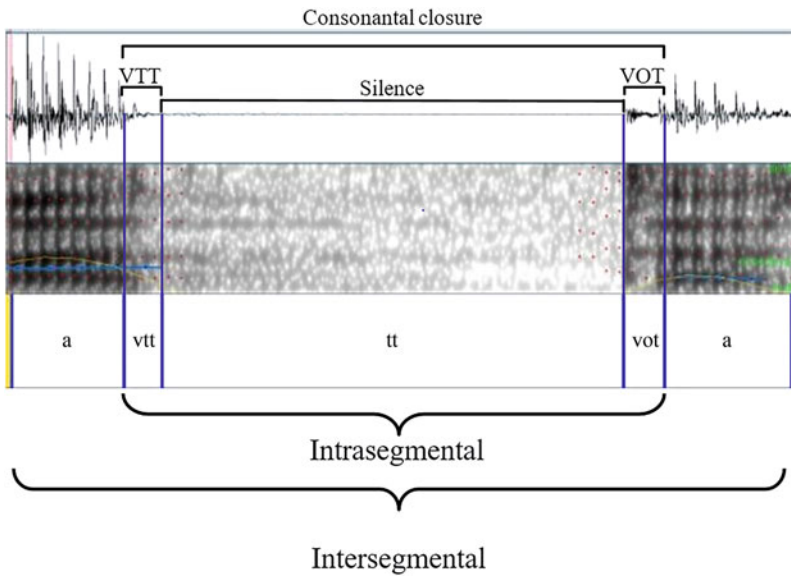


Fig. 1.1 Example of measurements of inter- and intrasegmental parameters in the acoustic signal of voiceless stops

number of samples for Taqbaylit. The differential significance of the compared pairs is given by the p , the t (size of the difference relative to the variation in sample data), and the df (degree of freedom) values. Only results with a probability of less than 5% ($p < 0.05$) were considered significant.

Results

Intersegmental Results for Voiceless Stops

The data for the three languages reveal that consonantal closure in absolute values is significantly longer for geminate stops than for their singleton counterparts (see Tables 1.2, 1.3 and 1.4).

The flanking vowels do not reveal any significant differences for singletons versus geminates in Tarifit and Taqbaylit. In Japanese, however, vowels (V1) are significantly longer before geminates than they are before singletons. Durational differences between geminates and singletons are maintained in fast speech despite compression of consonantal and vocalic segments (see Figs. 1.3, 1.4, and 1.5 below).

Table 1.2 Detailed T-test results of intersegmental data for Tarifit Berber in normal (NR) and fast (FR) speaking rates for voiceless stops

Interseg values (NR)	p value	t	df	Interseg values (FR)	p value	t	df
t vs tt	0.0001	11.13	5	t vs tt	0.0005	7.854	5
k vs kk	0.0009	7.109	5	k vs kk	0.0033	5.257	5
q vs qq	0.0001	10.54	5	q vs qq	0.0004	8.259	5

Table 1.3 Detailed T-test of intersegmental data for Taqbaylit Berber in normal (NR) and fast (FR) speaking rates for voiceless stops

Interseg values (NR)	p value	t	df	Interseg values (FR)	p value	t	df
t vs tt	0.0435	4.638	1	t vs tt	0.0344	5.255	1
k vs kk	0.0437	4.627	1	k vs kk	0.0237	6.377	1

Table 1.4 Detailed T-test results of intersegmental data for Japanese in normal (NR) and fast (FR) speaking rates for voiceless stops

Interseg values (NR)	p value	t	df	Interseg values (FR)	p value	t	df
p vs pp	0.0013	11.72	3	p vs pp	0.0059	7.039	3
t vs tt	0.0010	13.06	3	t vs tt	0.0015	11.35	3
k vs kk	0.0020	10.23	3	k vs kk	0.0024	9.622	3

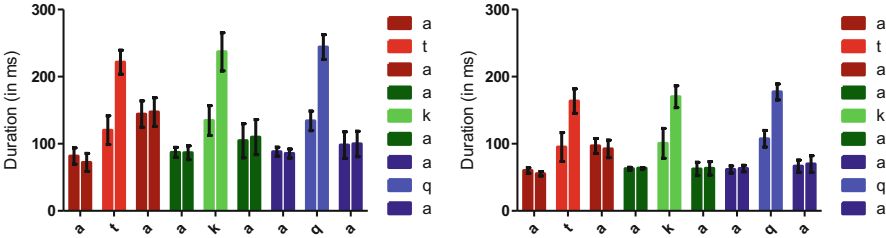


Fig. 1.3 Duration (ms) of intersegmental parameters of singleton (left bars) and geminate (right bars) voiceless stops in intervocalic positions, at normal (left) and fast (right) speaking rates for Tarifit

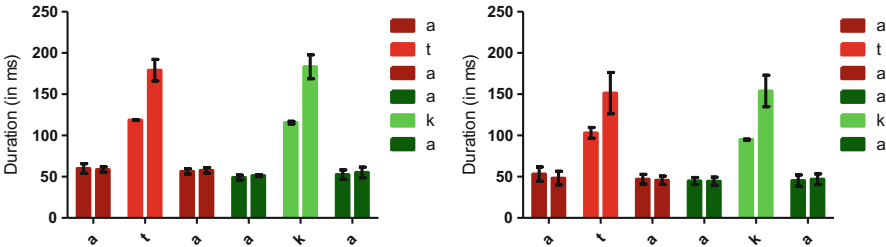


Fig. 1.4 Duration (ms) of intersegmental parameters of singleton (left bars) and geminate (right bars) voiceless stops in intervocalic positions, at normal (left) and fast (right) speaking rates for Taqbaylit

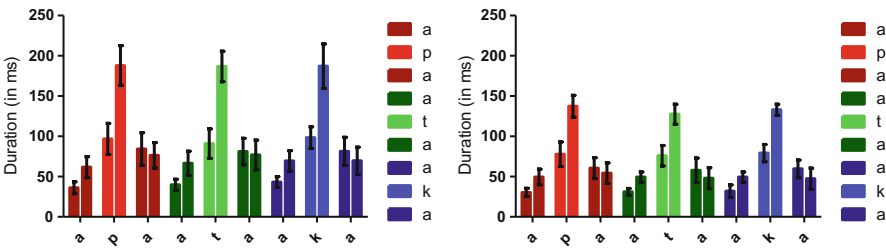


Fig. 1.5 Duration (ms) of intersegmental parameters of singleton (left bars) and geminate (right bars) voiceless stops in intervocalic positions at normal (left) and fast (right) speaking rates for Japanese. (Note that V1 is always significantly longer before geminates than before singletons)

Intersegmental Results for Voiced Stops

As in the abovementioned voiceless context, the data for the three languages reveal that consonantal closure, in absolute values, is significantly longer for geminate stops than for their singleton counterparts (see Tables 1.5, 1.6 and 1.7):

Table 1.5 Detailed T-test results of intersegmental data for Tarifit Berber in normal (NR) and fast (FR) speaking rates for voiced stops

Interseg values (NR)	<i>p</i> value	<i>t</i>	<i>df</i>	Interseg values (FR)	<i>p</i> value	<i>t</i>	<i>df</i>
d vs dd	< 0.0001	35.42	5	d vs dd	< 0.0001	30.97	5
g vs gg	< 0.0001	11.54	5	g vs gg	< 0.0001	21.03	5

Table 1.6 Detailed T-test results of intersegmental data for Taqbaylit Berber in normal (NR) and fast (FR) speaking rates for voiced stops

Interseg values (NR)	<i>p</i> value	<i>t</i>	<i>Df</i>	Interseg values (FR)	<i>p</i> value	<i>t</i>	<i>df</i>
d vs dd	0.0057	13.20	1	d vs dd	0.0149	8.091	1
g vs gg	0.0014	26.34	1	g vs gg	0.0034	17.14	1

Table 1.7 Detailed T-test results of intersegmental data for Japanese in normal (NR) and fast (FR) speaking rates for voiced stops

Interseg values (NR)	<i>p</i> value	<i>t</i>	<i>Df</i>	Interseg values (FR)	<i>p</i> value	<i>t</i>	<i>df</i>
b vs bb	0,0010	12.99	3	b vs bb	0,0010	13.00	3
d vs dd	0,0006	15.39	3	d vs dd	0,0008	13.94	3
g vs gg	< 0,0001	44.10	3	g vs gg	0,0031	8.815	3

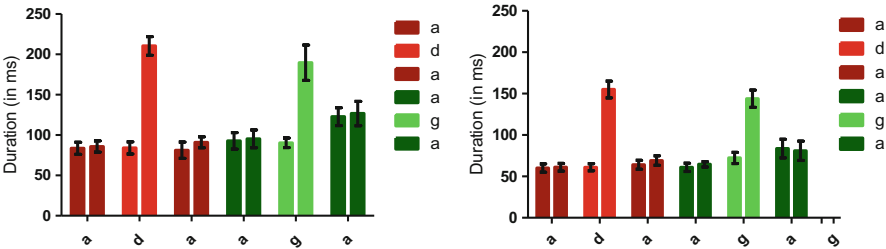


Fig. 1.6 Duration (ms) of intersegmental parameters of singleton (left bars) and geminate (right bars) voiced stops in intervocalic positions at normal (left) and fast (right) speaking rates for Tarifit

Here, also, the flanking vowels do not reveal any significant differences for singletons versus geminates in Tarifit and Taqbaylit. In Japanese, vowels (V1) are significantly longer before geminates than before singletons. For all three languages, durational differences between geminates and singletons are maintained in fast speech, notwithstanding compression of consonantal and vocalic segments (see Figs. 1.6, 1.7 and 1.8 below).

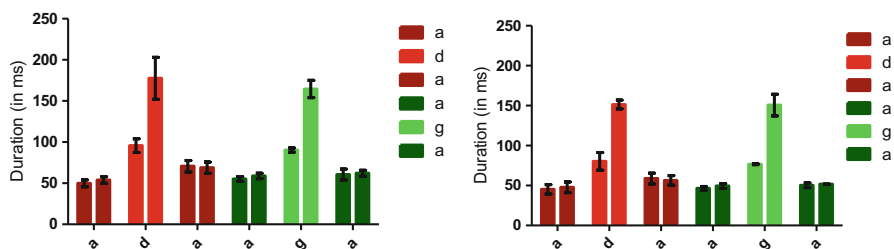


Fig. 1.7 Duration (ms) of intersegmental parameters of singleton (left bars) and geminate (right bars) voiced stops in intervocalic positions at normal (left) and fast (right) speaking rates for Taqbaylit

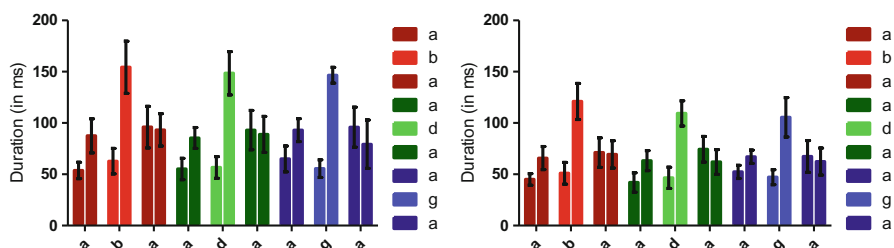


Fig. 1.8 Duration (ms) of intersegmental parameters of singleton (left bars) and geminate (right bars) voiced stops in intervocalic positions at normal (left) and fast (right) speaking rates for Japanese. (V1 is always significantly longer before geminates, than before singletons)

Intrasegmental Results for Voiceless Stops

In general, the acoustic silent phase of voiceless geminate stops is significantly longer than that of their singleton counterparts for all three languages in normal and fast speech regardless of segmental compression induced by increasing speaking rate (see Tables 1.8, 1.9 and 1.10).

VTT and VOT values do not differ between singletons and geminates (see Figs. 1.9, 1.10, and 1.11 below).

Intrasegmental Results for Voiced Stops

In this voiced context, the following observations apply to all three languages investigated: closure or occlusion of geminate stops is significantly longer than that of their singleton counterparts in both speaking conditions (see Tables 1.11, 1.12 and 1.13).

However, VOT values do not differ, like in the voiceless context, from one phonological category to the other (see Figs. 1.12, 1.13, and 1.14 below).

Table 1.8 Detailed T-test results of intrasegmental data for Tarifit Berber in normal (NR) and fast (FR) speaking rates for voiceless stops

Intraseg values (NR)	<i>p</i> value	<i>t</i>	<i>Df</i>	Intraseg values (FR)	<i>p</i> value	<i>t</i>	<i>df</i>
t vs tt-silence	< 0.0001	12.03	5	t vs tt-silence	0.0006	7.744	5
k vs kk-silence	0.0001	7.289	5	k vs kk-silence	0.0001	8.626	5
q vs qq-silence	0.0001	10.64	5	q vs qq-silence	0.0001	10.59	5

Table 1.9 Detailed T-test results of intrasegmental data for Taqbaylit Berber in normal (NR) and fast (FR) speaking rates for voiceless stops

Intraseg values (NR)	<i>p</i> value	<i>t</i>	<i>Df</i>	Intraseg values (FR)	<i>p</i> value	<i>t</i>	<i>df</i>
t vs tt-silence	0.0406	4.811	1	t vs tt-silence	0,0344	5.255	1
k vs kk-silence	0.0411	4.781	1	k vs kk-silence	0,0214	6.729	1

Table 1.10 Detailed T-test results of intrasegmental data for Japanese in normal (NR) and fast (FR) speaking rates for voiceless stops

Intraseg values (NR)	<i>p</i> value	<i>t</i>	<i>Df</i>	Intraseg values (FR)	<i>p</i> value	<i>t</i>	<i>df</i>
p vs pp-silence	0.0015	11.28	3	p vs pp-silence	0.0062	6.923	3
t vs tt-silence	0.0010	12.74	3	t vs tt-silence	0.0011	12.71	3
k vs kk-silence	0.0024	9.596	3	k vs kk-silence	0.0016	10.98	3

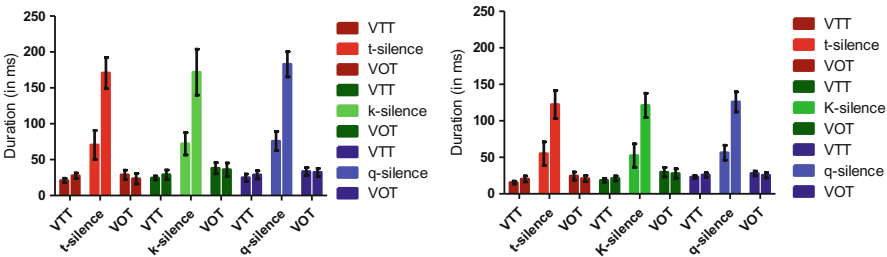


Fig. 1.9 Duration (ms) of intrasegmental parameters of singleton (left bars) and geminate (right bars) voiceless stops in intervocalic positions, at normal (left) and fast (right) speaking rates for Tarifit

Relative Values for Singletons Versus Geminates

Figures 1.15, 1.16, and 1.17 show typical examples of normalized consonantal data for apical consonants. Results are structurally similarly for all the other consonant types studied. The percentage of time taken by the target singleton or geminate consonant within the CV syllable allows observing trajectories (i.e., compression of

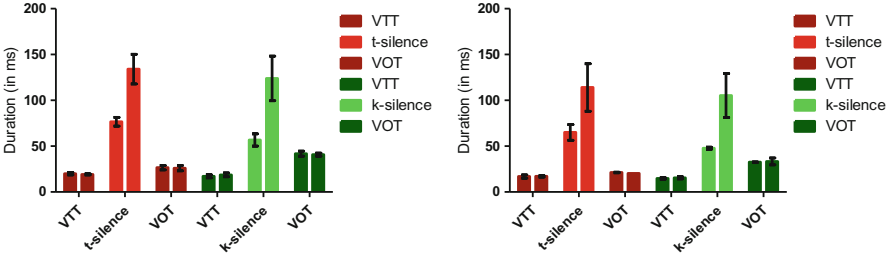


Fig. 1.10 Duration (ms) of intrasegmental parameters of singleton (left bars) and geminate (right bars) voiceless stops in intervocalic positions, at normal (left) and fast (right) speaking rates for Taqbaylit

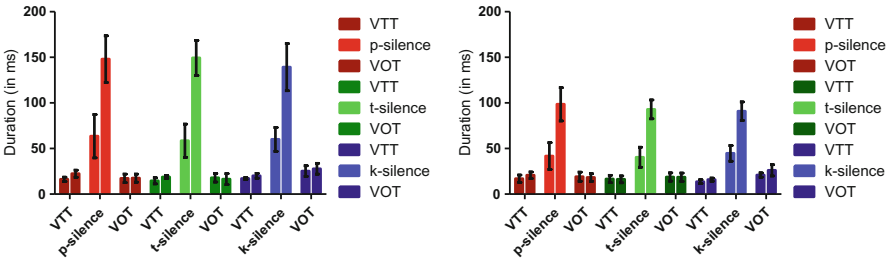


Fig. 1.11 Duration (ms) of intrasegmental parameters of singleton (left bars) and geminate (right bars) voiceless stops in intervocalic positions, at normal (left) and fast (right) speaking rates for Japanese

Table 1.11 Detailed T-test results of intrasegmental data for Tarifit Berber in normal (NR) and fast (FR) speaking rates for voiced stops

Intraseg values (NR)	<i>p</i> value	<i>t</i>	<i>Df</i>	Intraseg values (FR)	<i>p</i> value	<i>t</i>	<i>df</i>
d vs dd-occlusion	< 0.0001	33.20	5	d vs dd-occlusion	< 0.0001	26.29	5
g vs gg-occlusion	< 0.0001	12.07	5	g vs gg-occlusion	< 0.0001	20.16	5

Table 1.12 Detailed T-test results of intrasegmental data for Taqbaylit Berber in normal (NR) and fast (FR) speaking rates for voiced stops

Intraseg values (NR)	<i>p</i> value	<i>t</i>	<i>Df</i>	Intraseg values (FR)	<i>p</i> value	<i>t</i>	<i>df</i>
d vs dd-occlusion	0.0294	5.700	1	d vs dd-occlusion	0.0164	7.702	1
g vs gg-occlusion	0.0162	7.751	1	g vs gg-occlusion	0.0237	6.379	1

absolute CV values and increase, or decrease, of consonantal closure proportions within the CV span) followed by these phonological classes, as speaking rate varies. Such a projection in this two-dimensional space also contributes to neutralizing individual differences, which may be linked specifically to idiosyncratic differences in speaking rate. It can be seen from the data and in these figures that:

Table 1.13 Detailed T-test results of intrasegmental data for Japanese in normal (NR) and fast (FR) speaking rates for voiced stops

Intraseg values (NR)	<i>p</i> value	<i>t</i>	<i>Df</i>	Intraseg Values (FR)	<i>p</i> value	<i>t</i>	<i>df</i>
b vs bb-occlusion	0.0010	12.89	3	b vs bb-occlusion	0.0019	10.40	3
d vs dd-occlusion	0.0007	14.35	3	d vs dd-occlusion	0.0006	14.96	3
g vs gg-occlusion	0.0014	11.40	3	g vs gg-occlusion	0.0049	7.487	3

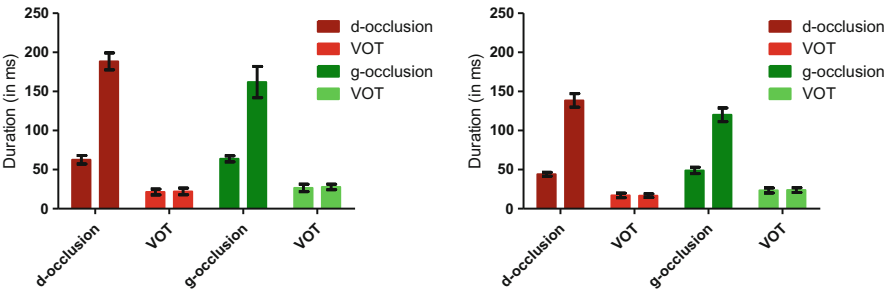


Fig. 1.12 Duration (ms) of intrasegmental parameters of singleton (left bars) and geminate (right bars) voiced stops in intervocalic positions, at normal (left) and fast (right) speaking rates for Tarifit

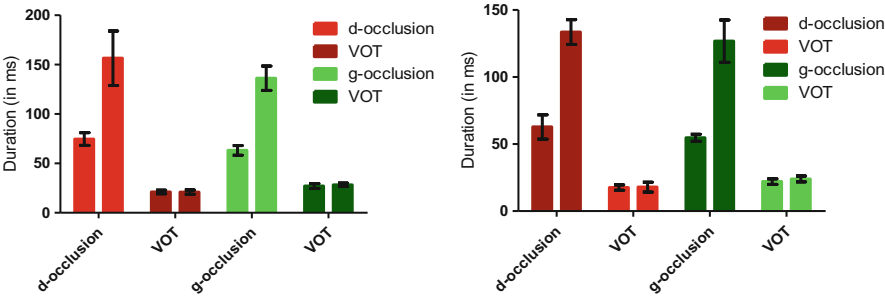


Fig. 1.13 Duration (ms) of intrasegmental parameters of singleton (left bars) and geminate (right bars) voiced stops in intervocalic positions, at normal (left) and fast (right) speaking rates for Taqbaylit

- 1) Phonological contrasts, well assured in relative terms for all speakers in normal speech, as well as in fast speech, are observable at the level of the consonantal phase (*y* axes), with a difference of about 10 to 20%, on average, depending on the language.
- 2) The distinction of classes is also made at the level of CV syllable durations (*x* axes). Typical examples for apical voiceless stops show, on average, in normal speech, a difference of 123 ms between singletons and geminates for Tarifit Berber, a difference of 75 ms for Taqbaylit Berber, and of 80 ms for Japanese.

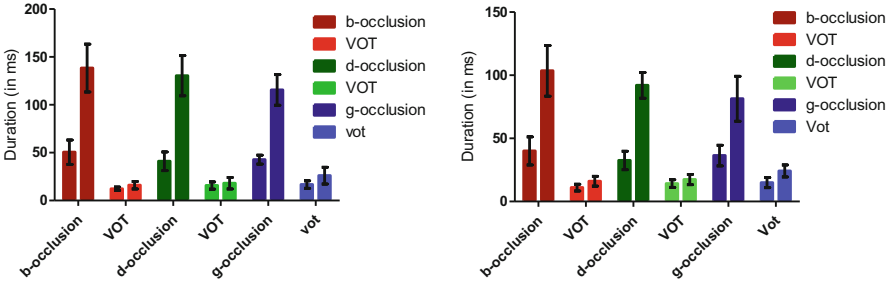


Fig. 1.14 Duration (ms) of intrasegmental parameters of singleton (left bars) and geminate (right bars) voiced stops in intervocalic positions, at normal (left) and fast (right) speaking rates for Japanese

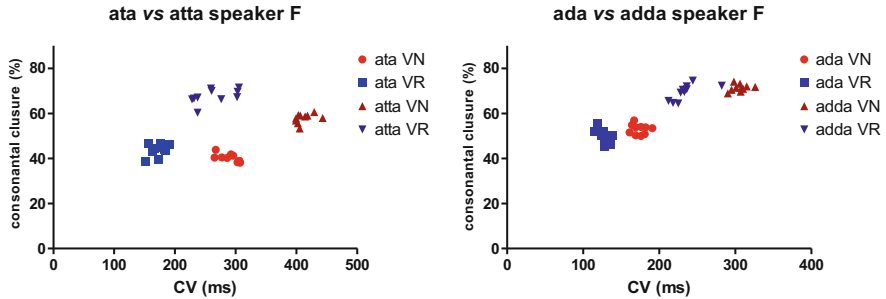


Fig. 1.15 The proportion of C/CV for /t/ vs /t/ (left) and /d/ vs /dd/ (right) in normal and fast speaking rates for Tarift Berber (Speaker F)

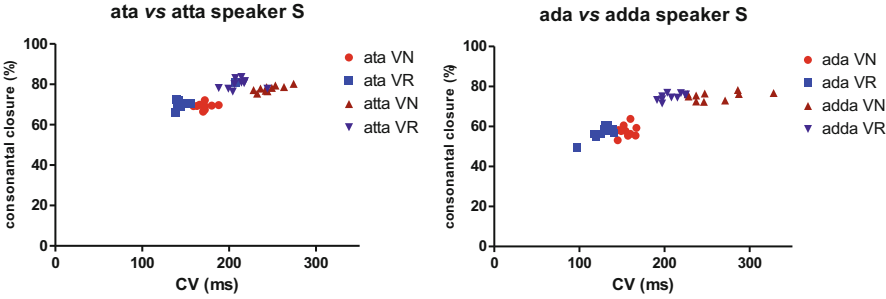


Fig. 1.16 The proportion of C/CV for /t/ vs /t/ (left) and /d/ vs /dd/ (right) in normal and fast speaking rates for Taqbaylit Berber (Speaker S)

3) With speech rate increase, CV syllables are compressed, but phonological categories remain separate despite this compression: a difference of 103 ms is measured in fast speech for Tarift Berber, one of 61 ms in Taqbaylit Berber, and a difference of 28 ms for Japanese. These values are like those observed in the voiced context. It seems that the phonological is less robust in Japanese.

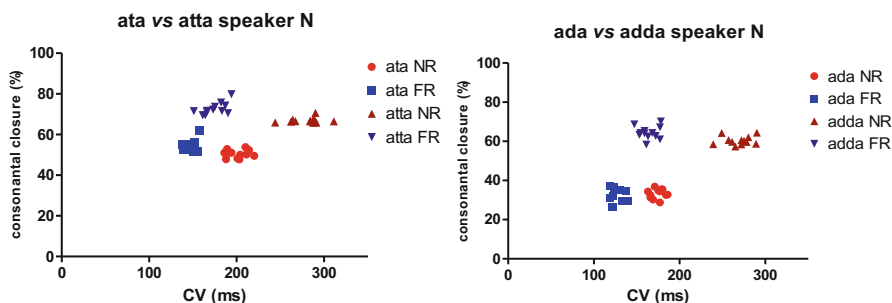


Fig. 1.17 The proportion of C/CV for /t/ vs /t:/ (left) and /d/ vs /dd/ (right) in normal and fast speaking rates for Japanese (Speaker N)

- 4) Strategies deployed by all speakers in the three languages, in order to preserve the phonological gemination contrast, can be summarized as follows: (a) when geminates reduce the duration of the CV syllable, singletons do as much, so as to somehow “avoid” any collision between the two categories; (b) by switching from normal speech to fast speech, the proportion of the consonantal closure for geminates and singletons may, either increase, or remain relatively stable. In all cases, strategies adopted by all speakers, in the three languages, serve to preserve phonological distinctiveness. This point will be taken up later.

Discussion and Conclusions

Absolute Durational Data

Results obtained in the intersegmental domain show, for all singleton and geminate consonants produced in normal or fast speech, that the absolute duration of consonantal closure is the main cue that distinguishes the two categories (in accordance with Hypothesis 1). This result corroborates those attested in the literature (see, e.g., (Abramson 1987, 1998; Bouarourou 2014; Bouarourou, et al. 2018a, b; Bouarourou, et al. 2010; Lahiri and Hankamer 1988; Local Simpson 1999; Ridouane 2003; Ridouane 2007). The duration of adjacent vowels does not contribute to contrasting gemination in Berber. However, the duration of the vowel, which precedes the target consonant does so in Japanese (see, e.g., (Campbell 1999; Fukui 1987; Hirata 2007; Hirose and Ashby 2007; Hussain and Shinohara 2019; Idemaru and Guion 2008; Itô and Armin 1995; Kawahara 2006; Kawahara 2015; Takeyasu 2012), as it is longer before geminates, than before singletons (Hypothesis 2 is, thus, only partially verified).

In the intrasegmental domain, analyses have led to the conclusion that the duration of the acoustic silent phase is the main correlate that distinguishes voiceless singleton plosives from their geminate counterparts (Hypothesis 3 is verified). As regards voiced singleton and geminate plosives, it is the duration of the occlusion

that serves to distinguish the two categories (this is consistent with Hypothesis 4). VOT is not a relevant candidate in separating the two phonological categories (this goes against Hypothesis 5). The same is true of VTT, which does not allow separating simple voiceless plosives from their geminate counterparts (Hypothesis 6 is not confirmed). Compression of segments, caused by increased speaking rate, does not interfere with preservation of phonological contrasts (Hypothesis 7 is, thus, confirmed). The fact that, for Japanese, the vowel is always significantly longer before geminates than before singletons in all consonantal contexts and in both speech conditions, contrary to findings for the two dialects of Berber, seems to be a language-specific factor (partially confirming Hypothesis 8: see *infra*, however). Let us recall also that results for Taqbaylit were obtained from quite a low number of participants. Hence, findings, in this particular investigation, must be taken with caution.

Relative Durational Data

This investigation has allowed depicting control of two temporal phenomena by analyzing intersegmental and intrasegmental relative values. Firstly, results confirm that absolute durations of the parameters observed are not more linked to speaker idiosyncratic speech rate differences than to the experimentally required variation of speech rate. Indeed, our relative timing analyses have had the advantage of normalizing the temporal data and, thus, erasing eventual individual temporal effects. Secondly, we have seen examples of preservation of the distinctiveness of phonological categories on the temporal dimension, even when the executional condition became difficult, namely, when speaking rate is increased. In addition, it was possible to underpin the various strategies that speakers adopt to preserve class distinctions in gemination. Indeed, these strategies, deployed to preserve phonological contrasts, rely, for singletons as for geminates, on:

- 1) Relative stability (Fig. 1.18) of the intersegmental consonantal phase, as a function of the CV syllable), while speech rate is increased. With increased speaking rate, while absolute duration of the CV syllable is compressed (along the x-axis), the percentage of time taken by the consonantal phase within this CV syllable remains relatively stable for both singletons and geminates.
- 2) A concomitant relative increase (Fig. 1.19), or reduction (Fig. 1.20) of the percentage of intersegmental (and intrasegmental) parameters as a function of the CV syllable. In these two cases, also, absolute duration of the CV syllable is compressed, while the proportion of time taken by the consonantal phase in this CV span is, either increased (Fig. 1.19), or reduced (Fig. 1.20).
- 3) Compression of CV and C intervals can also serve to maintain phonological class distinctions; when geminates are compressed, singletons also reduce the duration of the CV, or C intervals.

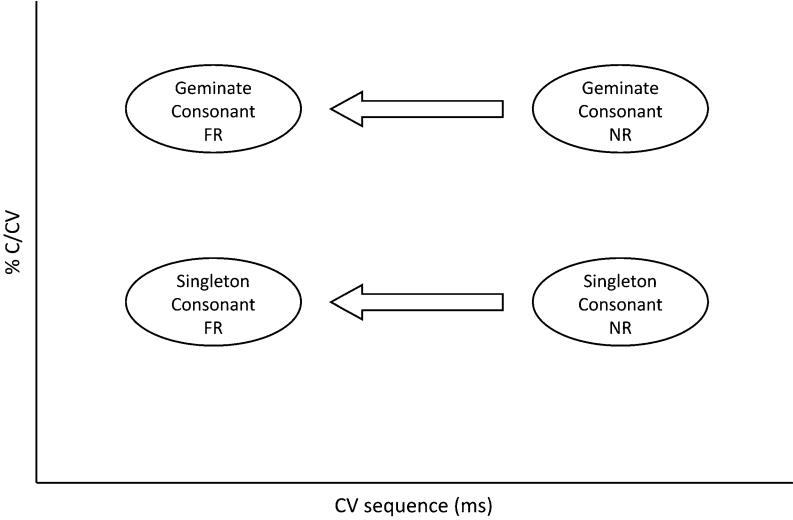


Fig. 1.18 Relative stability of the intersegmental consonantal phase as a function of the CV syllable with speech rate increase. NR Normal speech Rate, FR Fast speech Rate

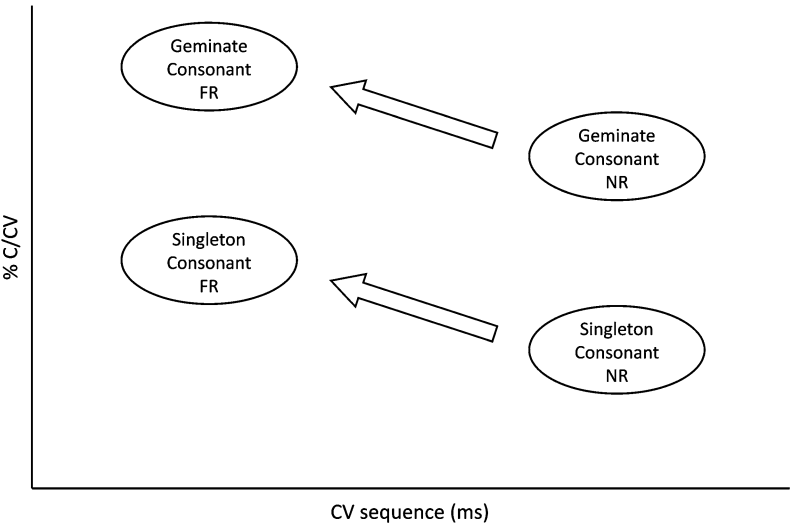


Fig. 1.19 Relative increase of the intersegmental consonantal phase as a function of the CV syllable with speech rate increase. NR Normal speech Rate, FR Fast speech Rate

Concluding Remarks

These results, observed in the CV domain, indicate that relative stability of intersegmental and intrasegmental consonantal parameters serves, cross-

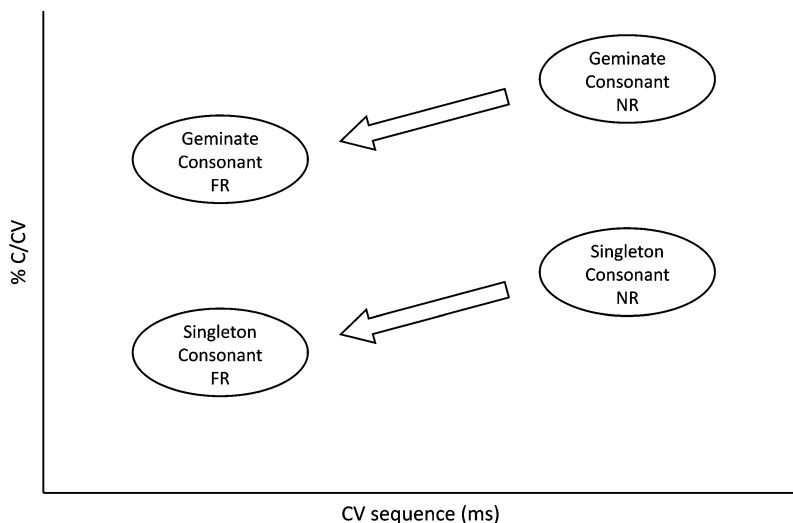


Fig. 1.20 Relative reduction of the intersegmental consonantal phase as a function of the CV syllable with speech rate increase. NR Normal speech Rate, FR Fast speech Rate

linguistically, to reinforce the phonological gemination contrast. Indeed, the proportion of time taken by intersegmental and intrasegmental consonantal parameters of geminates is systematically higher than that taken by their singleton counterparts, regardless of the speech rate condition. Note, that in case of variability of the proportion of intersegmental and intrasegmental consonantal parameters, maintaining the distinctiveness of the phonological categories is, nevertheless, based cross-linguistically also on a difference in the proportions of these parameters, with trajectories heading, so the speak, in parallel – in the same direction, in order to avoid any confusion as pertains to phonological classes. It should be mentioned here that, as concerns relative timing of singletons and geminates within the CV syllable, acoustic timing patterns for Japanese are like those observed for Tarifit and Taqbaylit Berber (this is not in accordance with Hypothesis 8).

We posit that the behavior these classes adopt is conditioned primarily by linguistic viability constraints; then, secondarily, by constraints related to the physical system of the phonatory apparatus [32]. Relative stability (see Fig. 1.18) as well as variability (see Figs. 1.19 and 1.20) do contribute to preserving phonological contrasts. In the case of relative stability, contrasts are maintained from normal to fast speech, solely along the consonantal phase; whereas, in the case of variability, phonological distinctions depend on, both consonantal phase, and CV durational differences.

Acknowledgments This work was partly supported by the Slovak Research and Development Agency under contracts No. APVV-15-0307.

References

- Abramson AS (1987) Word-initial consonant length in Pattani Malay. In: Proceedings of the 11th international congress of phonetic sciences 6. Academy of Sciences of the Estonian S.S.R., Tallinn, pp 143–147
- Abramson AS (1998) The complex acoustic output of a single articulatory gesture: Pattani Malay word-initial consonant length. In: Warotamasikkhadit U, Panakul T (eds) Papers from the fourth annual meeting of the southeast Asian linguistics society 1994. Arizona State University Program for SE Studies, Tempe, pp 1–20
- Agnello J (1975) Voice onset and voice termination features of stutterers. In: Furst LC, Webster LM (eds) Vocal tract dynamics and dysfluency. Speech and Hearing Institute, New York, pp 940–954
- Bouarourou F (2014) La gémation en tarifit: considérations phonologiques, *études acoustique et articulatoire*. PhD dissertation, University of Strasbourg, Strasbourg
- Bouarourou F, Vaxelaire B, Ridouane R, Hirsch F, Sock R (2010) La résistivité de la gémation en tarifit. In: Actes des 28e Journées d'Étude sur la Parole, Mons, pp 341–345
- Bouarourou F, Koya T, Bouzidi S, Vaxelaire B, Sock R (2018a) Cross-language and language-specific acoustic correlates of Geminata in Berber and Japanese. In: Proceedings of the 2nd international conference on natural language and Speech processing. ICNLSP, Algiers, pp 1–6
- Bouarourou F, Vaxelaire B, Laprie Y, Ridouane R, Sock R (2018b) The timing of geminate consonants in Tarifit Berber. *ELSEVIER Procedia Comput Sci* 128:25–31
- Campanell N (1999) A study of Japanese speech timing from the syllable perspective. *Onsei Kenkyuu* 3(2):29–39
- Fukui S (1987) Nihongo heisaon-no encho/tanshuku-niyoru sokuon/hisokuon toshite-no chooshu. *Onsei Gakkai Kaihou* 59:9–12
- Gaitenby J (1965) The elastic word. Haskins Laboratories, Status Report, Speech Research 2:1–12
- Ham WH (1998) Phonetic and phonological aspects of geminate timing. Ph.D. Dissertation, Cornell University
- Haraguchi S (2006) A theory of voicing. Meikai University, MS, Urays
- Hirata Y (2007) Durational variability and invariance in Japanese stop quantity distinction: roles of adjacent vowels. *Onsei Kenkyuu* 11(1):9–22
- Hirose A, Ashby M (2007) An acoustic study of devoicing of the geminate obstruents in Japanese. In: Trouvain J (ed) Proceedings of the XVIth international congress of phonetic sciences. Univ. des Saarlandes, Saarbrücken, pp 909–912
- Hussain Q, Shinohara S (2019) Partial devoicing of voiced geminate stops in Tokyo Japanese. *J Acoust Soc Am* 145:149–163
- Idemaru K, Guion SG (2008) Acoustic covariants of length contrast in Japanese stops. *J Int Phon Assoc* 38(2):167–186
- Itô J, Armin M (1995) Japanese phonology. In: Goldsmith J (ed) The handbook of phonological theory. Blackwell, Oxford, pp 817–838
- Jeannerod M (1988) The neural and behavioural organization of goal-directed movements, vol XII. Clarendon Press/Oxford University Press, New York
- Kawahara S (2006) A faithfulness ranking projected from a perceptibility scale: The case of voicing in Japanese. *Language* 82(3):536–574
- Kawahara S (2015) The phonetics of obstruent geminates, sokuon. In: Kubozono H (ed) The handbook of Japanese phonetics and phonology, Berlin, pp 43–78. Mouton [Handbooks of Japanese Languages and Linguistics 2]
- Klatt DH (1975) Voice onset time, frication, and aspiration in word-initial consonant clusters. *J Speech Hear Res* 18(4):686–706
- Koya T (2015) Etude de la gémation en japonai. Masters Dissertation, University of Strasbourg, France

- Lahiri A, Hankamer J (1988) The timing of geminate consonants. *Jornal of Phonetics* 16(3): 327–338
- Lehiste I (1970) *Suprasegmentals*. MIT Press, Cambridge
- Local J, Simpson AP (1999) Phonetic implementation of geminates in Malayalam nouns. In: Ohala JJ (ed) *Proceedings of the 14th international congress of phonetic science*. University of California Press, Berkeley, pp 595–598
- Ridouane R (2003) *Suites de consonnes en berbère: phonétique et phonologie*. PhD dissertation. Université Sorbonne-Nouvelle, Paris
- Ridouane R (2007) Gemination in Tashlhiyt Berber: an acoustic and articulatory study. *J Int Phon Assoc* 37(2):119–142
- Ridouane R (2010) Geminates at the junction of phonetics and phonology. *Laboratory Phonology* 10:61–90
- Smith JL (2006) Loan phonology is not all perception: evidence from Japanese loan doublets. In: Vance TJ, Jones KA (eds) *Japanese/Korean linguistics 14*. CSLI, Stanford, pp 63–74
- Sock R (1999) *Organisation temporelle en production de la parole. Emergence de catégories sensori-motrices phonétiques*. Presses Universitaires du Septentrion. Villeneuve d’Ascq, Darmstadt
- Sock R (2010) La quantité: implications pour un modèle de timing de la quantité vocalique. In: Choi-Jonin I, Duval M, Soutet O (eds) *Typologie et comparatisme. Hommages offerts à Alain Lemaréchal*, vol VIII–460. Peeters [Orbis Supplementa 28], Peeters Paris/Louvain, pp 413–430
- Sock R, Vaxelaire B (2009) How special is speech? In: Fuchs S, Loevenbruck H, Pape D, Perrier P (eds) *Some aspects of speech and the brain*. Peter Lang Internationaler Verlag der Wissenschaften, Berlin, pp 259–294
- Sock R, Vaxelaire B (2011) Spéciale, la parole! In: Huck D, Choremi T (eds) *Parole(s) et langue(s), espaces et temps*. Publications of the University of Strasbourg [Mélanges offerts à Arlette Bothorel-Witz], Starsbourg, pp 141–148
- Takeyasu H (2012) Sokuon-no chikaku-ni taisuru senkoo onsetsu shiin boin-no jizokujikanno eikyoo. *Onin Kenkyuu* 15:67–78

“Within-Category” Consonant–Vowel Interactions in Tashlhit: Labialized Consonant Alternations

2

Karim Bensoukas

Contents

Introduction	26
Facts and Previous Accounts	27
A Very Common Within-Category C–V Interaction: Labialized Consonant	
Dissimilation	28
A Very Rare Within-Category C–V Interaction: Labialized Consonant Assimilation	30
C ^w -Dissimilation As Identity Avoidance	34
C ^w -Assimilation As a Corollary of Vowel Copying	38
Vowel Copying	39
C ^w -Assimilation As a Case of Consonant Permeability	45
Is There a Commonality Between C ^w -Dissimilation and C ^w -Assimilation?	48
Tashlhit C–V Interactions and Conspiracy in Phonology	48
The Emergence of the Unmarked	49
Conclusion	50
References	51

Abstract

Tashlhit Berber presents two distinct, within-category C-V interactions involving the labialized velar and uvular consonants (C^w). First, these consonants undergo a dissimilation process mainly in the vicinity of the round vowel u- and the round glide w sporadically. This by now established dissimilatory phenomenon is a mundane, within-category C-V interaction. Second, labialized consonants also participate in a very intricate derounding phenomenon to be kept separate from dissimilation on account of the notable differences between the two. Offering an

I would like to thank Abdelaziz Boudlal, Nabila Louriz, Ayoub Noamane and an anonymous reviewer for their comments on this paper. Any errors of fact or analysis the final version of this paper may contain are my entire responsibility.

K. Bensoukas (✉)
Mohammed V University in Rabat, Rabat, Morocco
e-mail: k.bensoukas@um5r.ac.ma

Optimality-Theoretic analysis, the paper argues that while C^w -dissimilation, tautologically, involves a dissimilatory process, C^w -assimilation “derounds” labialized consonants as a side-effect of an assimilatory, vowel-copying operation. When entrapped in a vowel copying domain flanked by the agreeing unrounded vowels ([a . . a] or [i . . i]), labialized consonants deround in a rather assimilatory context. The latter process is an example of the impossible, or at best rare, within-category C-V interactions predicted by Padgett (2011: 1763).

Introduction

Consonant–vowel (C–V) interactions are a staple in overall treatments of features and phonological processes (e.g., Zsiga 2011; Morén-Duolljá 2011; Rose and Walker 2011). In a more detailed treatment, Padgett (2011) distinguishes “within-category” and “cross-category” C–V interactions. Unlike cross-category interactions, which are “C–V interactions in which the *primary* articulation of a [+consonantal] segment appears to interact with the place of a vowel,” “‘within-category’ interactions are those between a vowel and another (semi)-vocalic element, whether the latter is a secondary articulation or a primary one (a glide)” (Padgett 2011, pp. 1762–1764). Examples of within-category C–V interactions, with which I am concerned in this chapter, involve (i) vowels interacting with glides or secondary-articulated consonants, or conversely, (ii) consonants with adjacent vowels. Padgett reports on coda glides in Kabardian triggering the assimilation in the backness and rounding of the preceding vowels (e.g., /qəw/ → [quw] “swan”; /tsaj/ → [tsej] “one of wool (kind of coat)”; and on pharyngealized consonants in Ayt Seghrouchen Tamazight Berber, making the vowels back and lowered (e.g., [izi] “fly” and /izʕi/ → [ezʕe] “bladder”). Consonants are also affected by vowels as in Nupe labialization before rounded vowels (/ego/ → [egʷo] “grass”) and Russian palatalization before [i] or [e] (*stol* “table, Nom sg”/ *stolʲik* (Nom sg dim)/*stolʲik* (Loc sg). “It is worth noting that the examples of within-category assimilation . . . never involve a vowel changing the features of a glide or of a consonant’s secondary articulation. Such cases are at best rare, but it is not clear why they should be” (Padgett 2011, p. 1763).

In this chapter, I focus on two distinct, within-category C–V interactions involving the round dorsal (velar and uvular) consonants of Tashlhit, a variety of Amazigh (aka Berber) spoken in southwest Morocco. (Tashlhit is a member of the Berber family, now officially referred to in Morocco as Amazigh. Amazigh has morphological, lexical, and phonological affinities with Afroasiatic, of which it is a member. A relevant example is a rich consonantal system, which includes pharyngeals, uvulars, and glottals (Crass 2009, pp. 13–14). The geographical domain of Tashlhit is quite vast: the southern High Atlas, the Anti-Atlas, Bani, and the Sous plain, including major cities such as Agadir, Tiznit, and nearby cities (Boukous 1979, p. 13).) While other within-category C–V interactions exist in the language, emphasis spread being the most conspicuous, the alternations affecting labialized consonants are interesting in a twofold fashion. First, labialized consonants undergo a dissimilation process mainly in the vicinity of the round vowel *u*- and the round glide

w sporadically (Jebbour 1985; Elmedlaoui 1985, 1992/1995; Selkirk 1993; Bensoukas 2006, 2014). This by now established dissimilatory phenomenon is a mundane, within-category C–V interaction, as defined above. Second, labialized consonants also participate in a very intricate derounding phenomenon to be kept separate from dissimilation on account of the notable differences between the two, as will be shown in section “[Facts and Previous Accounts](#).” Henceforward, I will refer to these two C^w -alternations as C^w -dissimilation and C^w -assimilation, respectively, C^w standing for labialized consonants.

I will argue that while C^w -dissimilation, tautologically, involves a dissimilatory process, C^w -assimilation in contrast “derounds” C^w as a side effect of an assimilatory, vowel copying operation. After ample evidence is presented to sustain this distinction, a parallel Optimality-Theory (OT) account is provided (Prince and Smolensky 1993/2004; McCarthy and Prince 1993, 1995, 1999; see also Kager 1999; McCarthy 2002, 2007, 2008; and related works). The C^w -dissimilation analysis, essentially that in Bensoukas (2014), resorts to local constraint conjunction and stipulates that Tashlhit avoids the repetition of the feature [round] within the word/stem domain – a process analogous to the avoidance of labial consonants within the same domain. The C^w -assimilation analysis, on the other hand, shows that labialized consonants deround when entrapped in a vowel copying domain flanked by the agreeing unrounded vowels ([*a* . . *a*] or [*i* . . *i*]). By virtue of this entrapment, labialized consonants deround in a rather assimilatory context. This process is an example of the rare within-category C–V interactions quoted above.

While the chapter contributes to research on C–V interactions, it also contributes significantly to the elucidation of a long-standing, thorny issue in the phonology of Tashlhit. In comparison with previous analyses (Jebbour 1985; Elmedlaoui 1985; Lasri 1991), mine provides a unifying account of C^w -alternations, while at the same time reconciling the opposing processes of assimilation and dissimilation affecting a group of consonants in tandem, but not necessarily in a conspiracy fashion. The remainder of this chapter is organized as follows. Section “[Facts and Previous Accounts](#)” presents C^w -dissimilation and C^w -assimilation as well as surveys previous accounts. C^w -dissimilation and C^w -assimilation are analyzed separately in sections “[C^w-Dissimilation As Identity Avoidance](#)” and “[C^w-Assimilation As a Corollary of Vowel Copying](#),” respectively. Section “[Is There a Commonality Between C^w-Dissimilation and C^w-Assimilation?](#)” discusses possible commonalities between the two C–V interactions discussed in the chapter. Then, I conclude.

Facts and Previous Accounts

Tashlhit has five primary dorsal consonants, two velars and three uvulars, with labialization as a secondary articulation (represented above and henceforth as C^w): k^w , g^w , χ^w , ɸ^w , and q^w . (Tashlhit has three underlying vowels (*a*, *i*, *u*) and a larger gamut of consonants: The stops *b*, *t*, *d*, *tʰ*, *dʰ*, *k*, *g*, k^w , g^w , *q*, q^w ; the fricatives *f*, *s*, *z*, *sʰ*, *zʰ*, *ʃ*, *ʒ*, *ʃʰ*, *ʒʰ*, *χ*, *ɸ*, χ^w , ɸ^w , *ħ*, *ʕ*; the sonorants *m*, *n*, *l*, *r*, *lʰ*, *rʰ*; and the glides *w*, *j*. Practically, most of these have phonemic geminate counterparts. My transcriptions

conform to the IPA, and the data from the sources have been adapted to this convention.) Reference to these consonants as “labio-velar” or “round-dorsal” is not uncommon in the literature; however, the term I will use in this chapter is “labialized consonant.” I assume van de Weijer’s (2011) distinction between secondary articulation and double articulation involving the dorsum and lips. While the labialized consonants of Tashlhit (or the original KwaKwala example in van de Weijer) have secondary articulation, doubly articulated labial-velars involve structures that are equal, as in the Yetelenye (Papua New Guinea) and Eggon (a Bantu-Congo language of Nigeria) $k\bar{p}$ and $g\bar{b}$ sounds. Also, according to Ladefoged and Maddieson (1996, p. 356), “the addition of a lip rounding gesture is referred to as labialization. . . .”

The underlying labialized consonants of Tashlhit exhibit a very intriguing allophony resulting from two distinct within-category C–V interactions. The phonemic status of labialized consonants is established on the basis of the minimal pair test (Elmedlaoui 1985, pp. 172–173), syllable structure (Elmedlaoui 1985), and general phonetics, phonology, and morphology of Tashlhit (Lasri 1991, pp. 99–101). As far as the allophony is concerned, underlying labialized consonants surface phonetically as their corresponding non-round counterparts whenever the word they occur in contains the vocoid *u* or *w*. This C–V interaction dissimilatory process is one of the facets of a larger dissimilation process affecting primary labial consonants as well, a standard analysis which has achieved unanimity (Elmedlaoui 1985, 1992/1995; Jebbour 1985; Selkirk 1993; Bensoukas 2004; see Bensoukas 2014 for a comprehensive account), with the exception of Lasri (1991), which we will discuss in section “A Very Common Within-Category C–V Interaction: Labialized Consonant Dissimilation”.¹ There is one caveat, nonetheless: labialized consonants mysteriously deround in a non-dissimilatory context. This is the second C–V interaction involving labialized consonants. Regardless of the various attempts at explaining this derounding, it has resisted proper analysis. In the following subsections, I deal with each of these C–V interactions separately and review previous accounts.

A Very Common Within-Category C–V Interaction: Labialized Consonant Dissimilation

The first C–V interaction I present is that in which labialized consonants dissimilate in the vicinity of a rounded vocoid as in (1). (The details of the morphology of the items

¹Labial dissimilation, apparently an Afroasiatic phenomenon, affects Tashlhit agentive noun and reciprocal/reflexive labial affixes when they combine with roots containing a primary labial consonant (e.g., [anmgar] (Ag.N.) from underlying /a-m-mgar/ derived from *mgr* “harvest” and [nfassaj] (Ref./Recip.) from /m-fsj/ based on *fsi* “dissolve.” Round vowels or labialized consonants do not trigger this process. This dissimilation is similar to Barth’s law (Rubin 2008, p. 81), which turns a word-initial labial consonant *m* (followed by a vowel other than *u*) into *n* when the word contains another labial consonant, an example of which is Old Babylonian Akkadian [narkabtum] from underlying /markabtum/ “chariot”.

in (1) are not directly relevant to the discussion of C^w-dissimilation. Suffice it to say that the broken plural and preterite morphology relevant to this chapter generally alters the underlying vowel(s) of the base, creating contexts for C^w-alternations (concerning plurals, see Saib 1986; Jebbour 1988; and Idrissi 2000; in relation to preterites, see Basset 1929; Dell and Elmedlaoui 1991; Iazzi 1991, 2018; Bensoukas 2015).)

(1)	a.	<i>UR</i>	<i>Aorist</i>	<i>Preterite</i>	
	i.	/k ^w nu/	knu	k ^w ni/a	“bend”
		/g ^w nu/	gnu	g ^w ni/a	“sew”
		/χ ^w lu/	χlu	χ ^w li/a	“become crazy”
		/ɸ ^w lu/	ɸlu	ɸ ^w li/a	“be expensive”
		/q ^w lu/	qlu	q ^w li/a	“fry”
	ii.	/ak ^w r/	ak ^w r	ukr	“steal”
		/ak ^w z/	ak ^w z	ukz	“recognize”
		/ag ^w r/	ag ^w r	ugr	“be older”
		/ag ^w l/	ag ^w l	ugl	“hang”
		/ag ^w i/	ag ^w i	ugi	“refuse”
	b.	<i>UR</i>	<i>Singular</i>	<i>Plural</i>	
	i.	/ag ^w ru/	agru	ig ^w ra	“frog”
		/tag ^w lut/	taglut	tig ^w la	“oar”
		/aɸ ^w jjul/	aɸjjul	iɸ ^w jjal	“donkey”
		/aɸ ^w ruʃ/	aɸruʃ	iɸ ^w raʃ	“stick”
		/taɸ ^w rrust/	taɸrrust	tiɸ ^w rras	“fish-slice”
	ii.	/amddakk ^w l/	amddakk ^w l	imddukkal	“friend”
		/ag ^w l/	ag ^w l	uglan	“molar”
		/aɸ ^w i/	aɸ ^w i	uɸa	“calf”
		/aχ ^w s/	aχ ^w s	uχsan	“tooth”
		/aχχ ^w na/	aχχ ^w na	iχnwan	“buttock”

All the items in (1) contain underlying labialized consonants. In (1a-i), the aorist forms of the verbs all contain a final *u*, which induces the dissimilation of the labialized consonant, in contradistinction to the aorists in (1a-ii), in which the rounding of the dorsal consonant is maintained, absent the vowel *u*. The same behavior holds in the preterite in either case. The plural forms of the nouns in (1b) display exactly the same behavior. (The data on which the analysis is based comes from previous research done by the author. It has been collected in the region of Agadir, a Tashlhit-speaking area. The author is a native speaker of this variety of Amazigh. Additional data mentioned in the chapter comes from other varieties and is acknowledged in appropriate places in the chapter.)

The items in (1) are the first facet of the within-category C–V interactions analyzed in the present chapter. Tashlhit has a co-occurrence restriction on rounded sounds, namely, rounded vocoids and labialized consonants. What the data in (1) suggests is that whenever these two elements co-occur, a dissimilatory process is triggered. While it is possible to dissimilate the labialized consonant or the rounded vowel alike, Tashlhit opts for the dissimilation of the consonant. Recall

that this C^w-dissimilation is itself only one facet of the overall labial/round dissimilation in Tashlhit (see footnote 1).

In the literature on Tashlhit phonology, C^w-alternations have been approached from two different angles: They result from an “anchoring” process (Lasri 1991) or a “dissimilation” process (Jebbou 1985; Elmedlaoui 1985, 1992/1995; Selkirk 1993; Bensoukas 2006, 2014).

The anchoring analysis defends the idea that the rounding of Tashlhit labialized consonants is a vestige of the deletion of a vowel *u*. The analysis claims that a dorsal consonant is a suitable bearing unit for the rounding. If one exists in the stem, it saves the rounding of the vowel from disappearing, although the vowel itself does. For example, in a word like *knu* “bend,” the rounding of the final vowel is maintained when the vowel is altered or deleted and anchors to an available, suitable bearing element – a velar or uvular consonant; this would result in the form *kni*, which surfaces with C^w (*k^wni*), after anchoring the so to speak “floating” round feature on the dorsal consonant. This situation is analogous to the autosegmental stability behavior of tones when the vowels to which they are originally linked are deleted (Goldsmith (1976) and related works).

Conversely, the dissimilation analysis views C^w-dissimilation as a facet of the general dissimilation process of the language in which the primary labial consonants *b*, *f*, and *m* are involved, mutatis mutandis. A co-occurrence restriction that the language strictly observes dictates that the same stem should never contain two primary labial or two round specifications. Were two inimical specifications to co-occur within the stem domain, the one on the left would delete in the case of primary labials, as would the rounding of labialized consonants in the case of two round specifications.

The point is not so much to review the specifics of the anchoring analysis, nor to present the arguments against it, as it is to state my adherence to the essence of the dissimilation analysis as a working hypothesis. As the dissimilation account is quite established, I do not provide further elaboration at this point (see the references above for details and section “C^w-Dissimilation As Identity Avoidance” for a summary of the C^w-dissimilation in Bensoukas 2014).

A Very Rare Within-Category C–V Interaction: Labialized Consonant Assimilation

Labialized consonants appear without their rounding in other contexts, which *prima facie* suggests dissimilation. The first instance is that of passive verbs, as in (2b) (cf. the passive verbs not containing labialized consonants in (2a)). These can be described as involving the passive allomorph *ttj-* in addition to the epenthesis of the prefinal vowel *a* for the sake of augmenting the stem. (Two comments are due at this point. First, the passive allomorphs are [*ttu+*], [*tt+*], and [*ttj+*], the latter being used with vowel initial verbs. Detailed analyses of passive formation are available in Moktadir (1989) and Jebbou (1996). However, this allomorphy has not been addressed in the works mentioned, nor has the significant amount of variation passive morphology displays. An account of this aspect of passives is beyond the scope of the present chapter. Second, the prefinal vowel in passives, as well as

intensive aorists and action nouns, is treated as an epenthetic vowel resulting from augmentation (for this kind of approach, see Bensoukas 1994, 2001; Jebbour 1996). Stems need to conform to a prosodic requirement that makes their ultimate syllables heavy if they are not already so. This augmentation is blocked in forms that are underlyingly vowel final; unlike the passive verb forms in (2), which are all augmented by prefinal vowel epenthesis, a verb like /g^wnu/ “sew” in (3) has the corresponding passive form *ttg^wna*. As to the quality of the epenthetic vowel, it is either that of a default *a* or a copy of a base vowel. An anonymous reviewer raised the issue of not resorting in this case to schwa, the more common epenthetic vowel in Amazigh varieties. There are two reasons why schwa does not participate in augmentation: (i) schwa is not resorted to in Tashlhit for syllabification reasons, and the dialect relies heavily on syllabic consonants (see Dell and Elmedlaoui 1985, 2002; Boukous 1987, 2009, for example); (ii) even in dialects with schwa epenthesis, augmentation through schwa does not achieve the weight that augmentation achieves, schwa syllables being light. A detailed analysis of this variable syllable weight and related issues is available in Bensoukas (2006/2007, 2021).) The second instance is illustrated in (3), where action, agentive, and instrument nouns all contain a de-labialized consonant without there being a rounded vocoid in the vicinity. (The following abbreviations are used: Act. = action; Ag. = agentive; Aor. = aorist; C^w = labialized consonant; FCT = Feature Class Theory; Instr. = instrument; Int. = intensive; N. = noun; OCP = Obligatory Contour Principle; OT = Optimality Theory; Pass. = passive; Pl. = plural; Pret. = preterite; Recip. = reciprocal; Refl. = reflexive; RM = Realize Morpheme; Sg. = singular; TETU = the Emergence of the Unmarked; UR = underlying representation.)

- | | | | | |
|-----|----|---------------------|----------------------|---------------|
| (2) | | <i>UR</i> | <i>Passive</i> | |
| | a. | /amz ^s / | ttjamaz ^s | “catch” |
| | | /asj/ | ttjasaj | “take, carry” |
| | | /ass/ | ttjasas | “tie” |
| | b. | /ag ^w l/ | ttjagal | “hang” |
| | | /ak ^w r/ | ttjakar | “steal” |
| | | /ag ^w m/ | ttjagam | “draw water” |
| (3) | a. | <i>UR</i> | <i>Act.N.</i> | |
| | | /g ^w nu/ | tigni | “sew” |
| | | /k ^w nu/ | tikni | “bend” |
| | | /g ^w ru/ | tigri | “pick up” |
| | b. | <i>UR</i> | <i>Ag.N.</i> | |
| | | /g ^w nu/ | imgni | “sew” |
| | | /k ^w nu/ | imkni | “bend” |
| | | /g ^w ru/ | imgri | “pick up” |
| | | /k ^w ru/ | imkiri | “rent” |
| | c. | <i>UR</i> | <i>Instr.N.</i> | |
| | | /g ^w nu/ | issgni | “sew” |
| | d. | <i>Sg.</i> | <i>Pl.</i> | |
| | | tigmmi | tig ^w mma | “house” |

The ordered-rule analysis does not cover the entire data and is limited to cases similar to those in (3) only. In addition to resulting in an important generalization about the phonology of the language being missed, the absence of the passive data in (2) is an oversight that would require a wholesale revision of the rules in (4), and the entire analysis for that matter. (A reviewer added another objection to the ordered rules account in that the rule in (4.b) should not be referred to as a “vowel harmony rule,” which implies that vowels agree in rounding regardless of the value of this feature, but rather a “vowel derounding rule,” whereby a round vowel becomes unrounded, and not vice versa.)

The second analysis is the context disjunction analysis. In (5), I list the relevant contexts/rules, with *K* indicating round-velar consonants, *ob.* the compulsoriness of a rule, and *x* a singleton (as opposed to geminate) consonant:

(5) *Disjunction of C^w-assimilation contexts* (Elmedlaoui 1985, p. 185)

$$/K^w/ \rightarrow [K] \left/ \begin{array}{l} \left. \begin{array}{ll} \begin{array}{c} \text{(a)} \quad -(\dots) \overset{x}{\rule{0.5cm}{0.4pt}} a- \\ \text{(b)} \quad -(\dots) a \overset{x}{\rule{0.5cm}{0.4pt}} a (\dots)- \\ \text{(c)} \quad -(\dots) \overset{x}{\rule{0.5cm}{0.4pt}} (\dots) i =]_{n.fem.} \end{array} & \begin{array}{l} :ob. \\ :ob. \\ :ob. \end{array} \end{array} \right\} \end{array} \right.$$

A few comments, albeit non-inclusive, are in order. (5a) is a provision for the singular/plural pair *targa* “ditch” (*targ^wa)/*tirgg^win* (*tirgg^win) as well as the forms *ngg^wa* “cook, int. aor.” and *ttagg^wa* “look, int. aor.” (p. 179). (5b) accounts for C^w-assimilation in a pattern [*a* . . *a*] belonging to the same radical, as in the passive items listed in *ag^wl/ttjagal* “hang,” *ak^wz/ttjakaz* “recognize,” and *ak^wi/ttjakaj* “jump” (p. 181). Finally, (5c) accounts for C^w-assimilation in i-final feminine nouns that do not have a suffix *t*, as in *gru/g^wra/tigri* “pick up” versus *tak^wit* “calf, fem.”

In addition to all the details that lead to a significant loss of generalization, various objections can be expressed regarding the disjunction analysis. First, it is not evident how C^w-assimilation might ever be blocked by the phonological details of gemination. Second, limiting C^w-assimilation to feminine nouns resorts to morphological information that is not justified. Last but not least, the items under (5b–c) involve what is characterized in this chapter as vowel copying. If (5a) is excluded as a dialectal variation, which it is, the not so quite obvious link between C^w-assimilation and vowel copying becomes straightforward, and by the same token, most of the detail in the disjunction analysis becomes immaterial.

The third analysis explains C^w-assimilation as a result of vowel co-occurrence restrictions. Limiting the discussion to the C^w-assimilation in passive forms only, the analysis relies crucially on the co-occurrence restrictions on vowels in (6):

(6) *C^w-assimilation as vowel co-occurrence restrictions* (Lasri 1991, pp. 128–129):

Impossible vowel sequences:

u-i-u	i-u-i	a-u-a
u-a-u	i-a-i	a-i-a

In the co-occurrence restriction analysis, when two *a* vowels flank a labialized consonant, the latter is de-labialized as stated by Lasri (1991, p. 129):

La labialité des consonnes dorsales tombe lorsque deux voyelles identiques les cernent. Le seul déclencheur potentiel de cette délabialisation est l'occurrence de la suite /a-u-a/ (le 'u' faisant partie de la labialisée.)

Inherent in the co-occurrence analysis are a number of infelicities, only two of which I address. First, being restricted to passives only, the analysis misses a significant generalization, just like the disjunction rule analysis. Second, the vowel co-occurrence restrictions and their leading to de-labialization are belied by a number of facts outside passive morphology, as shown by masculine nouns like *ak^wfaj* “milk,” *ag^wmar* “horse,” *ak^wlal* “shell,” *alq^wnajn* “rabbit,” *aχ^wmaʒ* “scratch,” and feminine ones like *tag^wrsa* “plowshare” and *taχ^wda* “rod.”

To sum up, while there is consensus on the C^w-dissimilation approach to the Tashlhit C^w-alternations, C^w-assimilation has been approached in at least three different ways, none of which accounts for all the intricacies it presents. In the remainder of this chapter, I present an analysis of C^w-alternations as involving C^w-dissimilation in one case and C^w-assimilation in the other.

C^w-Dissimilation As Identity Avoidance

The analysis of C^w-dissimilation I propose is essentially that in Bensoukas (2014), in which the process is viewed as an identity avoidance effect (Yip 1995, 1998). (Various works treat identity avoidance (Alderete 1997; Suzuki 1998; Fukazawa 1999; Yip 1995, 1998; Itô and Mester 2003; Alderete and Frisch 2007; Bye 2011; Kawahara and Sano 2014 inter alia). Some of these consider it an OCP effect, the banning of two identical melodic elements (e.g., McCarthy 1986). In this chapter, I use the constraint conjunction approach (Smolensky 1993, 1995, 1997; Alderete 1997; Itô and Mester 2003), according to which one or more constraints can join forces in a specified domain. In the labialized consonant dissimilation case, I use the self-conjunction of the markedness constraint *Round applying in the stem domain.) Assuming a general tendency in the language to avoid the repetition of the feature [round], I explain two related aspects of the phenomenon: Dissimilation targets the labialized consonant and never the vowel, and it applies both from right-to-left and left-to-right. (A reviewer asks what would happen in forms like *CWu* and *WCu* if they existed. These forms do exist, with a minor difference, for example *xwu* “empty” and *ufu* “be more than enough.” A few other cases exist in which an otherwise expected alteration of an initial vowel to *u* results in dissimilation to *i* when the following sound is *w*; compare the aorist/preterite correspondents *arm/urm* “try” and *aws/iws* “help,” for example. While accounting for these two cases would certainly provide a more complete analysis of the phenomenon, as the reviewer suggested, it will definitely take us too far afield. Suffice it to say that dissimilation generally targets the secondary labialization of consonants and affects vocalic rounding only in a handful of forms in which the rounding that is affected is that of the “affixal” material.) The analysis is based on the self-conjunction of the markedness constraint *Round in the stem (*Rnd²_{Stem}), which bans two round specifications within that domain (see Smolensky 1993, 1995, 1997; Alderete 1997; Itô and Mester 2003). The constraint as well as the faithfulness

constraints it interacts with are listed in (7). The faithfulness constraint (7b-c) is discussed further below. An additional aspect of the analysis is to split the markedness constraint *Round into two more specific constraints formulated in (8a), inherently ranked with respect to one another as in (8b) (The claim is essentially that round consonants are more marked than round (semi-)vowels. In OT terms, a consonant with rounding incurs a more serious markedness violation than does a (semi-)vowel with rounding. The inherent ranking in (8b) is justified phonetically, since round velar consonants are complex segments with secondary articulation, as well as phonologically, given that Tashlhit labialized consonants lose their rounding to round (semi-)vowels):

- (7) a. *Rnd²_{Stem}: Two independent round specifications in a stem are banned.
 b. MAX(Rnd): An input rounding specification must have a correspondent in the output.
 c. MAX-Root(Rnd): An input rounding specification in the root must have a correspondent in the output.
- (8) a. Markedness constraints on the feature [round]:
 *Crnd: Round specifications in consonants are banned.
 *Vrnd: Round specifications in vocoids are banned.
 b. Ranking: *Crnd » *Vrnd

The various constraint interactions in C^w-dissimilation and the established individual rankings are as follows. First, since output forms avoid two round specifications by dropping one, *Rnd²_{Stem} dominates MAX(Rnd). Second, two faithfulness constraints MAX(Rnd) and MAX-Root(Rnd) are used to indicate the fact that labialized consonants are a characteristic of roots rather than affixes (I am using MAX-Feature, following Lombardi (1998)). To obtain this effect, MAX-Root(Rnd) dominates *Crnd, which in turn dominates MAX(Rnd), much in the spirit of Root-Faithfulness (McCarthy and Prince 1995; Beckman 1998). Finally, both MAX(Rnd) and MAX-Rt(Rnd) dominate *Vrnd, allowing for vowel rounding in both roots and affixes. One remaining ranking is that of *Rnd²_{Stem} and MAX-Rt(Rnd). I suggest establishing a dominance relation between *Rnd²_{Stem} and MAX-Rt(Rnd) to allow for the dissimilation of root rounding. (9) shows the round dissimilation constraint hierarchy established so far:

- (9) *Rnd²_{Stem} » MAX-Rt(Rnd) » *Crnd » MAX-Rnd » *Vrnd

To illustrate how the hierarchy in (9) works, I will discuss items to which morphology has applied minimally, and which are accordingly closest to their lexical representations. Then, I discuss items in which morphology has brought about a vowel change directly affecting C^w-dissimilation. In the latter case, an additional constraint on morpheme realization will be used.

The first case of C^w-dissimilation is that in items whose morphology is the simplest. An example is the input form /k^wnu/ “bend” containing two round

specifications and its corresponding aorist form. The Generator function of OT provides various candidates, the most informative of which are assessed in (10) (Conflict between faithfulness and markedness (and other) constraints is an inherent property of OT. For every input, the Generator function provides candidates to be assessed by the Evaluator function for output well-formedness against a constraint hierarchy. The evaluation is presented in a tableau, like (10), in which only informative candidates are included. The constraint to the left dominates the one to its right. In the text, we write *constraint*₁ » *constraint*₂ and separate the two constraints in tableaux by a solid line. If the constraints are not ranked with respect to one another, the two constraints are separated by a comma in the text and by a dotted line in tableaux. The output candidates are listed under the input, itself occupying the leftmost, top cell. The candidates are evaluated in an inclusive and parallel fashion against the constraint hierarchy. A star indicates a constraint violation; an exclamation mark following the star shows fatal violation. In principle, the candidate satisfying the higher ranked constraint(s) wins (is optimal) irrespective of the violation mark(s) it incurs for the lower ranked constraint(s), (a) violation (s) considered minimal. “☞” points at the optimal candidate):

(10)

/k ^w nu, Aor./	*Rnd ² _{Stem}	MAX-Rt(Rnd)	*Crnd	MAX-Rnd	*Vrnd
i. k ^w nu	*!		*		*
ii. kni		*!*		**	
iii. k ^w ni		*	*!	*	
☞ iv. knu		*		*	*

Candidate (10i) is the most faithful to the input, yet it incurs a fatal violation of identity avoidance. (10ii) drops both rounding specifications and fatally fails at root faithfulness. By keeping only one round specification, candidates (10iii) and (10iv) equally satisfy the constraint *Rnd²_{Stem} and tie, accordingly. The tie is resolved on the basis of markedness: with its consonant rounding, (10iii) is more marked than (10iv), which maintains vowel rounding. As I will show in detail in section “[The Emergence of the Unmarked](#)”, an Emergence of the Unmarked effect (TETU) (McCarthy and Prince 1994, 2004; Becker and Flack Potts 2011) can be identified in this case.

A similar point can be made regarding the singular form *agru* “frog” in tableau (11):

(11)

/ag ^w ru, Sg./	*Rnd ² _{Stem}	MAX-Rt(Rnd)	*Crnd	MAX-Rnd	*Vrnd
i. ag ^w ru	*!		*		*
ii. agri		*!*		**	
iii. ag ^w ri		*	*!	*	
☞ iv. agru		*		*	*

A contender I have not included is the candidate which maintains the secondary articulation and at the same time dissimilates it. It would be realized as $k^l nu$ in (10) and $ag^l ru$ in (11). (I would like to thank an anonymous reviewer for having brought this detail to my attention.) In a previous analysis, I pointed out that “one way we can explain this behaviour is the presence in the grammar of the language of a markedness constraint $*C^j$, which is obeyed even in contexts of vowel copying” (Bensoukas 2014, p. 168). Further evidence for this analysis can be provided on the basis of external databases, like *Phoible* (Moran and McCloy 2019). A quick search of the “Segments” tab in *Phoible* reveals that palatalized consonants, unlike labialized ones, are not attested in any of the Amazigh dialects reported on. The immediate consequence of this is that $*C^j$ is higher ranking, a status that is revealed in two different ways: (i) $*C^j$ dominates faithfulness constraints in relation to the inventory so that palatalized consonants are not possible phonemes of Amazigh; and (ii) $*C^j$ transcends inventory considerations and militates against any realization of palatalized consonants.

The second case of C^w -dissimilation is the one resulting from vowel ablaut in preterite and plural formation, changing an unrounded vowel into a rounded one or the opposite. The details put aside, the relevant constraint is Realize Morpheme (RM) (Kurusu 2001, p. 39). RM dominates the faithfulness constraint MAX-Rt(F) to ensure that the vowel ablaut process alters root vowels:

(12) *Realize Morpheme (RM)*:

Let α be a morphological form, β be a morphosyntactic category, and $F(\alpha)$ be the phonological form from which $F(\alpha+\beta)$ is derived to express a morphosyntactic category β . Then RM is satisfied with respect to β iff $F(\alpha+\beta) \neq F(\alpha)$ phonologically.

Two opposite cases, as it were, will be examined: The verb $/ag^w l/$, whose preterite is formed by ablauting the initial vowel to u , and the verb $/k^w nu/$, whose preterite is obtained by ablauting the final vowel to i/a , depending on person.

(13) *Dissimilated preterites*:

$/ag^w l, \text{ Pret.}/$	RM (Pret.)	$*Rnd^2_{\text{Stem}}$	MAX-Rt (Rnd)	$*C_{\text{rnd}}$	MAX- Rnd	$*V_{\text{rnd}}$
i. $ag^w l$	$*!$			*		
ii. $ug^w l$		$*!$		*		*
iii. $ug l$			*		*	*

Candidate (13i) fails to ablaut and fatally violates RM. (13ii) satisfies RM; nevertheless, by maintaining the input rounding specification of the labialized consonant, it violates $*Rnd^2_{\text{Stem}}$. (13iii) emerges as the optimal candidate, as it both satisfies RM and avoids two round specifications. (The dissimilating candidate $ig^w l$ is not included in (13). Although this candidate satisfies both $*Rnd^2_{\text{Stem}}$ and MAX-Rt(Rnd), it violates RM because it fails to realize the vowel u in the preterite.)

In (14), I discuss the opposite situation, viz., that of preterite verbs with a final vowel u ablaut. With the vowel u changing, these verbs remain with only one round

specification. This makes $*\text{Rnd}^2_{\text{Stem}}$ inactive and faithfulness decisive. Violating RM, candidates (14i-ii) are out of the competition, favoring (14iii) and (14iv) thus far. (14iv) loses because it drops the rounding specification of the labialized consonant for no reason, failing at faithfulness:

(14) *Non-dissimilated preterites:*

/k ^w nu, Pret./	RM (Pret.)	$*\text{Rnd}^2_{\text{Stem}}$	MAX-Rt (Rnd)	*Crnd	MAX-Rnd	*Vrnd
i. k ^w nu	*!	*		*		*
ii. knu	*!		*		*	*
iii. k ^w ni/a			*	*	*	
iv. kni/a			*!*		**	

C^w-dissimilation in broken plural forms is analogous to that in preterites, as shown in tableaux (15a-b), to which the comments on tableaux (13) and (14) apply to the letter:

(15) *a. Dissimilated plurals:*

/ax ^w s, Pl./	RM (Pl.)	$*\text{Rnd}^2_{\text{Stem}}$	MAX-Rt (Rnd)	*Crnd	MAX-Rnd	*Vrnd
i. ax ^w san	*!			*		
ii. ux ^w san		*!		*		*
iii. uxsan			*		*	*

b. Non-dissimilated plurals:

/ag ^w ru, Pl./	RM (Pl.)	$*\text{Rnd}^2_{\text{Stem}}$	MAX-Rt (Rnd)	*Crnd	MAX-Rnd	*Vrnd
i. ag ^w ru	*!	*		*		*
ii. igra			*!		*	
iii. ig ^w ra				*	*	

To summarize, the first within-category C–V interaction involving Tashlhit labialized consonants is a dissimilatory process in which the labialized consonant loses its rounding specification to a co-occurring stem vowel or glide. The OT analysis presented so far is in line with the standard dissimilation account of the phenomenon, be it linear (Jebbour 1985; Elmedlaoui 1985) or nonlinear (Elmedlaoui 1992/1995; Selkirk 1993).

C^w-Assimilation As a Corollary of Vowel Copying

The second within-category C–V interaction in Tashlhit is the C^w-assimilation process. Above, I have shown that this derounding occurs in vowel copy domains; in other words, it is the result of an assimilatory process. The aim of this section is to provide an OT analysis of C^w-assimilation, a prerequisite to which is the analysis of vowel copying, which I provide immediately.

Vowel Copying

Description and Analysis

In Tashlhit, both verb and noun forms display a vowel copy process, illustrated below with an epenthetic vowel and an affixal vowel. The analysis in this section is essentially that in Bensoukas (2014, 2017).

First, as illustrated in (16) and (17), when the verb base undergoes intensive aorist or action noun morphology, an augmentation process through vowel epenthesis takes place (Bensoukas 1994, 2001; Jebbour 1996). The prefinal epenthetic vowel defaults to *a* in consonant-only bases (16a and 17a) or copies the already existing base vowel (16b and 17b). (Various works point out the alternating quality of the prefinal vowel of Amazigh intensive aorists (Basset 1929; Abdel-Massih 1968; Iazzi 1991, 2018; Dell and Elmedlaoui 1991; Boukhris 1986; Bensoukas 2001, 2017, and a host of others). In Basset’s (1929: L-LI) words: “Pour ce qui est des alternances vocaliques fondamentales, la forme d’habitude reproduit exactement le vocalisme de l’aoriste. . . La voyelle est généralement de timbre a . . . Parfois de timbre a, i, ou u, elle n’est que le reflet de la voyelle de l’aoriste.”) In a quite simplified fashion, the initial vowel in the Act.Ns in (17a) is an inflectional affix, which does not appear in nouns derived from a vowel-initial base, as in (17b), claimed to be underlyingly *u*-initial (e.g., Iazzi 1991, 2018). (Issues related to the initial vowel in Amazigh nouns are complex, since they involve phonological, morphological, and syntactic aspects, and the accounts suggested remain quite controversial. One trend is to distinguish an affix initial vowel from a radical one, each of which can display a behavior of its own when the noun takes on different inflections such as number and case. The words *argaz* “man” and *argan* “argan trees/oil” look quite similar on the surface, but their Construct State forms are different: *urgaz* and *wargan*, respectively, with the former losing its inflectional affix vowel and the latter maintaining its radical one when Construct State *u* is added. A similar behavior is displayed by feminine nouns in their Construct State forms: the affix vowel is deleted, e.g., *tmkart*, sg./*tmkarin*, pl.; and the root vowel is maintained, e.g., *targant*, sg./*targinin*, pl.. Another complexity is seen in the different plural patterns in feminine nouns, as in the pairs *tamkart*/*timkarin* and *targant*/*targinin*, which show different initial vowels. Detailed presentations and analyses are to be found in Basset (1932, 1945), Saib (1982), Guerssel (1983), Jebbour (1988, 1991), Idrissi (2000, 2001), and Lahrouchi (2013) inter alia.) Other changes that do not have a bearing on the analyses in this chapter are the absence of *tt* in *s*-initial (causative or pseudo-causative) intensive aorists, the degemination of some intensive aorists, and the gemination in some action nouns:

- (16) a. *Default vowel epenthesis:*
- | | | |
|--------|---------|------------------|
| /fdʕr/ | ttfdʕar | “have breakfast” |
| /knkr/ | tknkar | “pick a bone” |
| /skr/ | skar | “do” |
- b. *Copy vowel epenthesis:*
- | | | |
|-----------|---------|---------------|
| i. /azzl/ | ttazzal | “run” |
| /kawl/ | tkkawal | “be in hurry” |
| /sawl/ | sawal | “speak” |

ii.	/bikks/	ttbikkis	“fasten one’s belt”
	/mmiz ^ɕ z ^ɕ g/	ttmiz ^ɕ z ^ɕ ig	“yawn”
	/ssird/	ssirid	“wash”
iii.	/mmutt/	ttmuttul	“mingle/intertwine”
	/ggunzr/	ttgunzur	“have a nosebleed”
	/ssumm/	ssumum	“suck”

(17) a. *Default vowel epenthesis:*

UR	Aor.	Act.N.	
/frn/	frn	afran	“sort out (grains)”
/rd ^ɕ l/	rd ^ɕ l	art ^ɕ t ^ɕ al	“lend”
/k ^w ti/	k ^w ti	ak ^w ttaj	“remember”

b. *Copy vowel epenthesis:*

/Uf ^ɕ /	ff ^ɕ	uf ^ɕ	“go out”
/Ukd ^ɕ /	kkd ^ɕ	ukud ^ɕ	“blind”
/Ug ^w z/	gg ^w z	uguz	“descend”
/Uk ^w m/	kk ^w m	ukum	“beat”

Second, the initial inflectional nominal vowel, like the one in (17a), is also subject to vowel copying in a set of underlyingly vowel-final verbs such as the *u*-final verbs in (1a-i) above and the morphologically related verbs in (18). In these verbs, the nominal forms appear with a final vowel *i*, and I assume that vowel to be part of the nominal morphology of this type of verbs.

(18)	UR	Aor.	Noun	
	/ɤra/	ɤr	tɤri	“read”
	/rz ^ɕ a/	rz ^ɕ	tirz ^ɕ i	“break”
	/rɤa/	rɤ	tirɤi	“be hot”
	/lsa/	ls	timlsit	“wear”
	/mmntra/	mmntra	timmntrit	“beg”

Three patterns of Amazigh vowel copying emerge in which the copying epenthetic or affixal vowel is a replica of the verb base vowel, illustrated in (19) with intensive aorists:

(19)	a.	[a . . a]:	/ɤawl/	ttɤawal	“run”
	b.	[i . . i]:	/mmiz ^ɕ z ^ɕ g/	ttmiz ^ɕ z ^ɕ ig	“yawn”
	c.	[u . . u]:	/mmutt/	ttmuttul	“entangle, refl”

The analysis of Tashlhit vowel copying facts rests on a few assumptions, in which I follow treatments of vowel harmony phenomena, those couched in OT included (Baković 2000; Beckman 1997, 1998; Gafos and Dye 2011; van der Hulst and van de Weijer 1995; Krämer 2003; Nevins 2010; Ní Chiosáin and Padgett 2001; Padgett

2002; Rose and Walker 2011; Walker 2012, among others). First is Feature Class Theory (FCT) (Padgett 2002), according to which features are members of various classes, and constraints on well-formedness may target individual features as well as classes of features. Second, vowel copying in FCT is viewed as spreading the whole vocalic constituent in an all-or-nothing fashion (see Clements and Hume 1995; van der Hulst and van de Weijer 1995, for example) and resulting in a feature-sharing structure. These two assumptions are detailed in (20), with *V* representing the epenthetic vowel and *b*, *h*, *r* standing, respectively, for *back*, *high*, and *round*:

(20) a. *Feature classes* (2002:96):

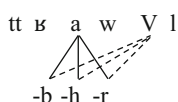
Vplace: {high, low, back, round...}

Height: {high, low}

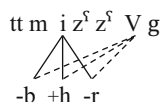
Color: {back, round}

b. *FCT approach to Tashlhit vowel copying*:

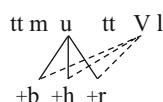
i. [a...a]



ii. [i...i]



iii. [u...u]



Third, constraint violation in vowel copying is assessed gradiently (e.g., Beckman 1997; Ní Chiosáin and Padgett 2001). In this view, linked structures like those in (20b) are associated with one specification and, by virtue of that, incur a single violation mark of the markedness constraint ($*aF$). Contrariwise, structures with similar or different feature specifications in a one-to-one fashion incur multiple violations of the markedness constraint(s) in question ($**aF$ or $*aF + *bF$).

Now, I present the analysis proper. The faithfulness constraint MAX-V(F), which is similar to the MAX constraint above, and the markedness constraints on individual vowels, given in shorthand as $*a$, $*i$, and $*u$, interact with the constraint AGREE and NOLINK, the constraint against linked structures adapted from Itô, Mester and Padgett (1995) (Considering vowel markedness as a scale of harmony (Prince and Smolensky 1993/2004), in which $[i]$ is less marked than $[u]$, and $[a]$ less marked than both $[i]$ and $[u]$, is convenient. The scale $a > i > u$ can be converted into the hierarchy $*u \gg *i \gg *a$, in which each markedness constraint is the sum of the individual violations each vowel incurs. For ease of exposition, I will be proceeding this way in the remainder of this chapter. Candidates can equally be assessed against detailed individual constraints (see Bensoukas 2017).):

(21) *Constraints on vowel copy*:

- a. MAX-V(F): An input vowel feature has a correspondent in the output.
- b. AGREE-V(F): Vowel features must agree.
- c. NOLINK: Features must not be multiply linked.

- d. *a: Assign a violation mark to each occurrence of *a*.
 *i: Assign a violation mark to each occurrence of *i*.
 *u: Assign a violation mark to each occurrence of *u*.

Illustration is provided with the $[a \dots a]$ and $[i \dots i]$ patterns. Crucial to the analysis proposed below is the idea that even the bases containing only vowels *a* involve vowel copying, albeit “vacuously.” (By “vacuous” I mean that the process of copying takes place even if the quality of the epenthetic vowel is akin to the default one. The representations in tableaux (22ii–iii) and (23ii–iii) show the difference between the defaulting candidates and copying ones. C^w-assimilation constitutes the strongest argument for this position.) The intensive aorist forms of the verbs *ḡawl* ‘be in a hurry’ and *mmiz^zg* ‘yawn’ are used as examples. The $[u \dots u]$ pattern will be commented on further below. Epenthetic quality will always seek to be the least marked structure, depending on what input vowel constituency is.

Take the $[a \dots a]$ pattern in (22), for instance. The unfaithful candidate (22i) is ruled out by virtue of its incurring a fatal violation of the high-ranking MAX-V(F). Candidate (22ii) epenthesises the vowel *i* and is accordingly more marked than a defaulting candidate- one with an epenthetic *a*. Candidates (22iii) and (22iv) both have an epenthetic *a*- a tie. With its one-to-one associations with the same vowel features, the defaulting candidate (22iii) incurs more violation marks than does the feature-sharing candidate (22iv). Although the latter incurs a violation of NOLINK, that violation is immaterial, and it is accordingly considered the optimal candidate. For space reasons, AGREE stands for AGREE-V(F) in the tableaux to come.

(22) The $[a \dots a]$ pattern

/ḡawl, Int.Aor./	AGREE	MAX-V(F)	*u	*i	*a	NOLINK
i. tt ḡ i w i l -b +h -r -b +h -r		*!		**		
ii. tt ḡ a w i l -b -h -r -b -r +h	*!			*!	*	
iii. tt ḡ a w a l -b -h -r -b -r -h					*!*	
iv. tt ḡ a w a l -b -h -r					*	*

The $[i \dots i]$ pattern illustrated in (23) follows more or less the same logic. The defaulting candidates (22i–ii) are out because of faithfulness and markedness

considerations, respectively. The copying candidate (23iv) incurs less markedness violations than the closest candidate (23iii), violation of NOLINK being once more immaterial.

(23) The [i...i] pattern

/mmiz ^h z ^h g, Int.Aor./	AGREE	MAX-V(F)	*u	*i	*a	NOLINK
i. <div>tt m i z^h z^h a g -b +h -r -b -h -r</div>	*!			*	*!	
ii. <div>tt m a z^h z^h a g -b -h -r</div>		*!			*	*
iii. <div>tt m i z^h z^h i g -b +h -r -b +h -r</div>				*!*		
iv- <div>tt m i z^h z^h i g -b +h -r</div>				*		*

In a nutshell, Tashlhit displays a vowel copying process in which all its three vowels participate, yielding the patterns [a . . a], [i . . i], and [u . . u]. The context is one in which a vowel is epenthesized to augment a stem or one which contains an underspecified affix vowel. The vowel copying process can be argued to compete as to markedness with the defaulting strategy. A TETU effect is to be noted with respect to vowel copying, to which I will return in some detail in section “[The Emergence of the Unmarked](#)” below.

Round Dissimilation and Round Copying: The Reconciliation

The multiple linking of the round feature as the one illustrated in (20b–iii) and tableau (24) poses a composite problem: (i) Why does it not qualify as a breach of identity avoidance, and (ii) why is it not used as a strategy to avoid identity?

The answer to the first question is simple: Linked structures, unlike one-to-one associations, incur one violation of markedness only, which puts them outside the purview of *Rnd²_{Stem}. Tableau (24) below shows this clearly. By dint of there being only one underlying round specification, the feature-sharing structure incurs a NoLink, rather than MAX-V(F), violation. In candidate (24i), the epenthetic vowel is realized by an independent round feature, which results in its violation of *Rnd²_{Stem}. Candidate (24ii), on the contrary, avoids identity by realizing the epenthetic vowel as a copy of the input one – a feature-sharing typical of spreading in assimilation contexts.

(24) *Tashlhit vowel copying: Feature-sharing*

/mmuttl, Int.Aor/ rnd	*Rnd ² _{Stem}	MAX- Rt(Rnd)	*Crnd	MAX- Rnd	*Vrnd	No Link
i- tt- m u t t u l rnd rnd	*!				**	
ii- tt- m u t t u l V rnd					*	*

The second question, why multiple linking is not used as a strategy to avoid identity, requires more elaboration. If multiple linking were to apply as a means of satisfying the identity avoidance constraint, the two offending rounding features would be fused into one which is doubly linked to two segments. Tashlhit does not resort to this option because it is costly. The two cases – vowel copying and the co-occurrence of two round specifications – are different at the input level: In the vowel copying case, there is only one round specification, whereas two underlying round segments come with two specifications. When two underlying round sounds end up sharing one feature, the loss of one input specification incurs a violation of MAX-(F), as illustrated in (25), where the notation <rnd> indicates failure to maximize an input round feature.

(25) *Tashlhit C^w-dissimilation: No feature-sharing*

/k ^w n u/ rnd rnd	*Rnd ² _{Stem}	MAX- Rt(Rnd)	*Crnd	MAX- Rnd	*Vrnd	No Link
i. k ^w n u rnd rnd	*!		*		*	
ii. k ^w n u <rnd> rnd		*	*!	*	*	*
iii. k n u <rnd> rnd		*		*	*	

The faithful candidate (25i) fails at identity avoidance. By eliminating one input round specification, candidates (25ii–iii) satisfy *Rnd²_{Stem}, through deletion and double linking in the former case and mere deletion in the latter. The tie between the two candidates is resolved by the lower markedness constraint, declaring suboptimal the more marked candidate – (25ii).

In a nutshell, it is the nature of the input – one round feature in copying versus two in dissimilation – that determines the crucial difference between round dissimilation and round copying. As shown in (24) and (25), the different outputs are shaped essentially by faithfulness.

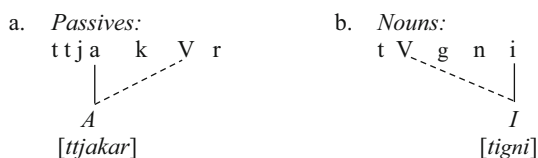
C^w-Assimilation As a Case of Consonant Permeability

The notion of “consonant permeability” is borrowed from Ni Chiosáin and Padgett (2001). The relevant facts we will consider are those relating to the derounding of labialized consonants in non C^w-dissimilation contexts, a reminder of which is in (26):

- (26) a. *Passives*: /ag^wl/ ttjagal “steal”
 b. *Nouns*: /g^wnu/ tigni (Act.N) issgni (Ag.N) “sew”
 c. *Sg./Pl. pair*: tigm̥mi tig^wmma “house”

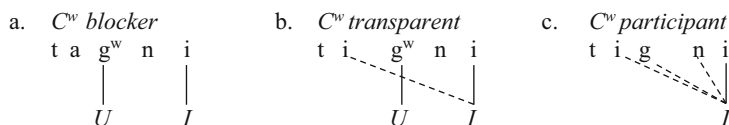
Tashlhit does definitely not allow the co-occurrence of two or more rounding specifications within a stem. Quite strikingly, however, the forms in (26) mysteriously deround labialized consonants, regardless of the absence of a round segment. More than just coincidentally, C^w-assimilation in (26), as has been shown above, takes place in domains that may be characterized as involving vowel copying. The representations in (27) are examples, where *I* and *A* stand for the features associated with the vowels *i* and *a*, respectively. Inkelas (1994) suggests that if an alternation is predictable, this is a case of underspecification. I follow this proposal and consider the affix vowel underspecified, hence its representation as *V*.

(27) *Total vowel copying in passive and nominal forms:*



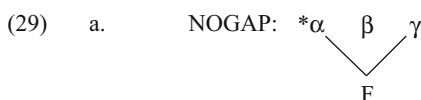
Locality considerations are quintessential in assimilatory cases, particularly those involving vowel copying/harmony. In this respect, the labialized consonant intervening between the trigger and target of copying acts in one of the following ways: (i) it is opaque and blocks the “spreading” of the vocalic features, (ii) it is transparent to the copying operation (Steriade 1987; Odden 1994; Ni Chiosáin and Padgett 2001; Nevins 2010; Zsiga 2011; Rose 2011; Gafos and Dye 2011; Walker 2012, among others). Another possibility is that (iii) it participates in it. These scenarios are represented in (28) using the nominal form *tigni*. Only (28c) is attested in Tashlhit either in the [*i* . . . *i*] or the [*a* . . . *a*] patterns.

(28) *Consonants in harmony domains:*



Scenario (28c) is inspired by the work of Ní Chiosáin and Padgett (2001), who claim that strict locality demands that all consonants intervening between the trigger of vowel harmony and its target be participants in the process. It is the notion of “consonant permeability” proposed by the authors that lies at the heart of (28c) and plays a crucial role in explaining Tashlhit C^w-assimilation.

The approach I defend in this chapter is that C^w-assimilation takes place when the consonant is entrapped in a vowel copying domain. Since the labialized consonants surface without their rounding, the constraint MAX-R(Rnd) is dominated. The fact that the labialized consonant participates in the copying process is an argument for the fact that NOGAP, a constraint demanding strict locality (see Itô et al. 1995, as well as Padgett 2002 and references mentioned therein), should outrank MAX-R(Rnd). The constraint and the ranking are provided in (29):



b. *NOGAP* » *MAX-R(Rnd)*:

/ak ^w r, Pass. /	NOGAP	MAX-R(Rnd)
i. ttjak ^w ar	*!	
☞ ii. ttjakar		*

Let’s illustrate the idea with the action noun *tigni*. Tableau (30) does not include the constraints *Rnd²_{Stem}, MAX-Rnd, and *Vrnd, these being irrelevant to the evaluation:

(30) *Consonant permeability in Tashlhit*:

/tV+g ^w nu/	AGREE	NOGAP	MAX-R(Rnd)	*Crnd	*i	*a
i. <i>Blocking</i> : $\begin{array}{ccccc} t & a & g^w & n & i \\ & & & & \\ A & U & I & & \end{array}$	*!			*	*	*
ii. <i>Transparency</i> : $\begin{array}{ccccc} t & i & g^w & n & i \\ & & \text{---} & & \\ & U & & I & \end{array}$		*!		*	*	
iii. <i>Participation</i> : $\begin{array}{ccccc} t & i & g & n & i \\ & & \text{---} & & \\ & & & I & \end{array}$			*		*	

Candidate (30i) illustrates the blocking effect. Spelling out the affixal vowel as the default vowel, this candidate fatally fails at vowel copying. Candidate (30ii), although satisfying vowel copying demands, fatally violates strict locality. Candidate (30iii) satisfies both vowel copying and strict locality. The higher order constraint requiring strict locality of vowel copying forces it to be unfaithful to the

rounding of the labialized consonant. Since this violation is minimal, (30iii) is optimal.

Passive forms are amenable to a similar analysis. Candidate (31i) spells the epenthetic vowel as the default vowel, which simply coincides with the quality of the root vowel. If the radical vowel were different, *i* for example, the difference in quality would be more apparent. Candidate (31i), the most faithful, violates the constraint AGREE and fails in the competition. The transparent candidate (31ii) violates the constraint NOGAP. The optimal candidate is (31iii), whose labialized consonant derounds as a result of its participating in the copying process.

(31)	/ak ^w r, Pass. /	AGREE	NOGAP	MAX-R(Rnd)	*Crnd	*a
i. <i>Blocking:</i>	ttjaik ^w a _r	*!			*	**
ii. <i>Transparency:</i>	ttjaik ^w a _r		*!		*	*
iii. <i>Participation:</i>	ttjaika _r			*		*

To wind the C^w-assimilation analysis up, a tentative explanation can now be offered to the isolated case of the sg./pl. pair *tigmmi*/*tig^wmma* “house,” in which the C^w-alternation is not motivated by identity avoidance. Since the presence of the labialized consonant *g^w* would not be otherwise predictable, one can only assume that it is part of the underlying representation of *tigmmi*. Two verbal bases can be postulated for the noun *tigmmi*. First, *gum* and *umu*, both with the meaning of “contain,” are attested in different varieties of Tashlhit. On the assumption that these evolved from some form like *g^wmu*, the underlying corresponding noun *tV-g^wmi* surfaces as *tigmmi*, in a similar fashion to the akin verb/noun pair *g^wnu*/*tigni*, whose noun has been shown to involve vowel copying. Second, a verb base *g^wmu* exists, meaning “grow up.” Once more, a similar account resting on C^w-assimilation due to participation in the copying process can be established. In this view, the word *tigmmi* would literally mean “something that contains” or “a place where one grows up,” respectively. In either case, C^w-assimilation applies according to the constraint hierarchy established so far, as shown in (32):

(32)	/tV+g ^w mu/	AGREE	NO GAP	MAX-R(Rnd)	*Crnd	*u	*i	*a
i.	tag ^w mmi	*!			*		*	*
ii.	tig ^w mmi		*!		*		*	
iii.	tigmmi			*			*	

The same comments under tableau (30) above can be reiterated here. The plural form *tig^wmma* does not involve a copying process, and as such does not require the derounding of the consonant *g^w*.

To recapitulate, the account we have proposed for recalcitrant C^w-assimilation in Tashlhit is based on the observation that labialized consonants entrapped in vowel

copying domains lose their rounding without there ever being the motivating identity avoidance factors. Labialized consonants deround since strict locality in vowel copying instigates consonants intervening between trigger and target to participate in the process – a consonant permeability effect.

Is There a Commonality Between C^w-Dissimilation and C^w-Assimilation?

To finish the analysis, it would be nice to seek a commonality between the two facets of C^w-alternations. Since C^w-dissimilation and C^w-assimilation achieve the same result, viz., dropping the round specification of the labialized consonant, the possibility of there being a conspiracy in phonology pops up first. In this section, I will claim that although a conspiracy proper cannot explain C^w-alternations, reduction of markedness through TETU can be pursued as a promising line of analysis.

Tashlhit C–V Interactions and Conspiracy in Phonology

That different rule applications “conspire” toward the achievement of the same surface phonological structure has been noted since early generative phonology. Kisseberth (1970) made the argument that rules conspire toward the achievement or avoidance of certain phonological structures on the basis of the phonology of Yawelmani, in which different vowel epenthesis and consonant deletion rules conspire to avoid what can be characterized as complex syllable margins. Revisiting the argument, the author shows different areas where conspiracies arise (Kisseberth 2011), reporting on the synchronic hiatus resolution in Chicano Spanish and the restriction on nasal-voiceless stop sequences in Bantu languages, which resort to more than one strategy. Kisseberth provides further evidence for conspiracy in phonology from Universal Grammar (e.g., the various strategies used in compliance with the OCP), language acquisition, and loanword phonology. Although indisputable, this “functional unity” of rules was difficult to formalize in the early generative models. With its output-oriented constraints and the different outcomes ensuing from constraint interaction, OT presents itself as a model in which conspiracy in phonology receives a straightforward account. In OT, McCarthy (2002, 2007) characterizes situations where different processes converge to achieve the same result, i.e., cases of conspiracy, as involving “homogeneity of target/heterogeneity of process.” The author provides an example of how OT handles conspiracy in phonology using hiatus resolution.

The argument in the present chapter is that C^w-alternations in Tashlhit fall out from the opposing processes of dissimilation and assimilation (vowel copy). The result in either case is the derounding of the labialized consonants. At first blush, this suggests a homogeneity of target-heterogeneity of process pattern. However, the rationale behind each process is quite different: While C^w-dissimilation achieves identity avoidance, in an OCP-like fashion, C^w-assimilation, being an assimilatory

process, leads to similarity between the segments involved. The two processes are actually the opposite of each other; accordingly, the “functional unity” that is a definitional criterion of conspiracy does not seem to be met. To use McCarthy’s terms, although there is heterogeneity of process, homogeneity of target seems to be missing.

In brief, it is difficult to sustain a strict conspiracy interpretation of the C^w-alternations in Tashlhit.

The Emergence of the Unmarked

Although conspiracy cannot strictly be used to account for the C^w-alternations, I will show that C^w-dissimilation and C^w-assimilation have more in common than meets the eye: Both processes result in the reduction of markedness. In so doing, I will be using the notion of the Emergence of the Unmarked (TETU) (McCarthy and Prince 1994, 2004; Becker and Flack Potts 2011). Unmarked structures typically emerge when a markedness constraint that is normally inactive because it is crucially dominated becomes active in contexts where the higher-ranked constraint is not relevant. I will show that TETU can be identified in both C^w-dissimilation and vowel copying.

In the C^w-dissimilation tableaux (10) and (11) above, the relevant portions of which I repeat in (33), the tie between the seriously competing candidates is resolved on the basis of the markedness constraint *Crnd, which is itself dominated. When all the relevant constraints are unable to decide on the optimality of the candidates, the dominated markedness constraint becomes decisive and leads to a TETU effect in dissimilation.

(33) TETU in Tashlhit C^w-dissimilation:

a.	/k ^w nu, Aor./	*Rnd ² _{Stem}	MAX-Rt(Rnd)	*Crnd
	i. k ^w ni		*	*!
	^{CP} ii. knu		*	

b.	/ag ^w ru, Sg./	*Rnd ² _{Stem}	MAX-Rt(Rnd)	*Crnd
	i. ag ^w ri		*	*!
	^{CP} ii. agru		*	

A similar TETU effect can be discerned in vowel copying, as shown by the relevant portions of tableaux (22) and (23) above, repeated in (34) for convenience. In (34a), the markedness constraint *a, which is otherwise crucially dominated in situations where faithfulness prevails, becomes active in epenthetic vowel copying situations in which faithfulness plays no role. In all three copying patterns, violations of *a are fatal, and it is actually one additional violation of this constraint that decides between the defaulting candidate and the less marked copying one in (34a). The TETU effect is even clearer in the [u...u] copying candidate in (34b). This approach rests on the assumption that the constraint AGREE can be done away with:

(34) *TETU in copying patterns:*

a.	/ɤawl, Int.Aor./	*u	*i	*a
i.				*!*
ii.				*
b.	/mmiz ^ɤ z ^ɤ g, Int.Aor./	*u	*i	*a
i.			*	*!
ii.			*	

The point is that in both C^w -dissimilation and vowel copying in Tashlhit, TETU effects are discernible. In (33), dissimilating the labialized consonant instead of the round vowel yields a less marked structure. In the copying process, the defaulting candidates incur one more violation of the markedness constraint $*a$ than the copying candidates. Since the constraint that decides is a markedness constraint that is otherwise inactive, this is a case of TETU in vowel copying. If the TETU analysis can be generalized to the entirety of the processes involved, then C^w -alternations are actually unified by an urge to reduce markedness.

Conclusion

This chapter has dealt with two related C–V interactions involving labialized consonants in Tashlhit. In both C^w -dissimilation and C^w -assimilation, C^w is affected by vowels in the vicinity, with the presence of u/w leading to identity avoidance and the application of a vowel copying process leading to consonant permeability. I have argued that while C^w -dissimilation accounts for one facet of the C^w -alternations in Tashlhit, the other facet is definitely not a dissimilatory process. Derounding is shown, in contrast, to be the corollary of an assimilatory, vowel copying operation – an example of the rare within-category C–V interactions. OT analyses of the various phonological phenomena related to the two C–V interactions involving labialized consonants have been established.

While the chapter contributes to research on C–V interactions, it is also a significant contribution to research on the phonology of the language insofar as it sheds light on a long-standing thorny issue in the phonology of Tashlhit. In comparison with previous accounts, the analysis in this chapter unifies labialized consonant alternations, while at the same time reconciling the opposing processes of

assimilation and dissimilation affecting a group of consonants in tandem. Although, the possibility of the C^w-alternations strictly ensuing from a conspiracy in phonology has been shown to be untenable, the two facts of these C-V interactions have been shown to have reduction of markedness as a common trait.

References

- Abdel-Massih ET (1968) Tamazight verb structure: a generative approach. Indiana University Publications, Bloomington
- Alderete J (1997) Dissimilation as local conjunction. *NELS* 27:17–32
- Alderete J, Frisch SA (2007) Dissimilation in grammar and the lexicon. In: de Lacy P (ed) *The Cambridge handbook of phonology*. Cambridge University Press, Cambridge, pp 379–398
- Baković E (2000) Harmony, dominance and control. PhD dissertation, Rutgers, The State University of New Jersey
- Basset A (1929) *La langue berbère. Morphologie. Le verbe- Etude de thèmes*. Leroux, Paris
- Basset A (1932) Sur l'état d'annexion en berbère. *Bull Soc Ling Paris* 33(2):173–174
- Basset A (1945) Sur la voyelle initiale en berbère. *Rev Afr* 1945:82–88
- Becker M, Flack Potts K (2011) The emergence of the unmarked. In: van Oostendorp M, Ewen CJ, Hume EV, Rice K (eds) *The Blackwell companion to phonology*, vol 3. Wiley-Blackwell, Malden/Oxford, UK, pp 1363–1378
- Beckman J (1997) Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology* 14:1–46
- Beckman J (1998) Positional faithfulness. PhD dissertation, University of Massachusetts, Amherst
- Bensoukas K (1994) Tashlhit agentive nouns- an optimality-theoretic approach. D.E.S. thesis, Mohammed V University, Faculty of Letters, Rabat
- Bensoukas K (2001) Stem forms in the nontemplatic morphology of Berber. Doctorat d'Etat thesis, Mohammed V University, Faculty of Letters, Rabat
- Bensoukas K (2004) Markedness, faithfulness and consonant place in Tashlhit roots and affixes. *Lang Lit* 18:115–153
- Bensoukas K (2006) The emergence of the unmarked in Tashlhit round velar consonant dissimilation. In: Allati A (ed) *La Linguistique amazighe: Les nouveaux horizons*. Publications of the Faculty of Letters, Tétouan, Tétouan, pp 76–118
- Bensoukas K (2006/2007) Variable syllable weight in Amazighe. *Lang Ling* 18/19:31–58
- Bensoukas K (2014) Featural dissimilation in Tashlhit- avoiding the repetition of [labial] and [round]. Publications of FLSH-Rabat, Rabat
- Bensoukas K (2015) Amazigh negative verb morphology- how far does the variation go? *Brill's J Afro-Asiatic Lang Ling* 7(2):211–278
- Bensoukas K (2017) Stem-augmentation in Amazigh intensive aorists: default vs. copy epenthesis. To appear in *Actes du premier colloque international de linguistique berbère- en hommage à André Basset*. March 9-10, 2017, LACNAD/INALCO- Paris
- Bensoukas K (2021) Schwa as a non-moraic vowel in Amazigh: an optimality-theoretic account. *Asinag* 16:61–90
- Boukhris F (1986) *Le verbe en tamazight: Lexique et morphologie*. (Parler des Zemmours). Third cycle thesis, Université Paris III- EPHE (4ème Section)
- Boukous A (1979) Le profil sociolinguistique du Maroc. *Bull écon soc du Maroc* 140:5–31
- Boukous A (1987) Phonotactique et domaines prosodiques en berbère. Doctorat d'Etat thesis, Université Paris VIII- Vincennes-Saint-Denis
- Boukous A (2009) *Phonologie de l'amazighe*. IRCAM Publications, Rabat
- Bye P (2011) Dissimilation. In: van Oostendorp M, Ewen CJ, Hume EV, Rice K (eds) *The Blackwell companion to phonology*, vol 3. Wiley-Blackwell, Malden/Oxford, UK, pp 1408–1433

- Clements GN, Hume EV (1995) The internal organization of speech sounds. In: Goldsmith J (ed) *The handbook of phonological theory*. Blackwell, Cambridge, MA, pp 245–306
- Crass J (2009) Afroasiatic languages. In: Brown K, Ogilvie S (eds) *Concise encyclopedia of the languages of the world*. Elsevier, Amsterdam, pp 12–15
- Dell F, Elmedlaoui M (1985) Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *J Afr Lang Ling* 7:105–130
- Dell F, Elmedlaoui M (1991) Clitic ordering, morphology and phonology in the verbal complex of Imdlawn Tashlhiyt Berber (Part II). *Lang Orient Anciennes- Philol Ling* 3:77–104
- Dell F, Elmedlaoui M (2002) Syllables in Tashlhiyt Berber and in Moroccan Arabic. Kluwer, Dordrecht
- Elmedlaoui M (1985) *Le parler chleuh d’Imdlawn (Maroc)- Ségments et syllabation*. Third Cycle thesis, Université Paris VIII, Vincennes-Saint-Denis
- Elmedlaoui M (1992/1995) *Aspects des représentations phonologiques dans certaines langues chamito-sémitiques*. PhD dissertation, Mohammed V University, Faculty of Letters and Human Sciences, Rabat. [Published by the Faculty of Letters and Human Sciences, Rabat]
- Fukazawa H (1999) *Theoretical implications of OCP effects on features in optimality theory*. PhD dissertation, University of Maryland at College Park
- Gafos AI, Dye A (2011) Vowel harmony: opaque and transparent vowels. In: van Oostendorp M, Ewen CJ, Hume EV, Rice K (eds) *The Blackwell companion to phonology*, vol 4. Wiley-Blackwell, Malden/Oxford, UK, pp 2194–2180
- Goldsmith J (1976) *Autosegmental phonology*. PhD dissertation, MIT, Cambridge, MA. [Distributed by IULC]
- Guerssel M (1983) A phonological analysis of the construct state in Berber. *Ling Anal* 11:309–330
- Iazzi E (1991) *Morphologie du verbe en tamazight- (Parler des Ait Attab, Haut Atlas Central)- Approche prosodique*. D.E.S. Thesis, Mohammed V University, Faculty of Letters, Rabat
- Iazzi E (2018) *Norme et variations en amazighe marocain (Aspects morpho-phonologiques)- Pour une approche polynomique de l’aménagement linguistique*, Doctorat d’Etat thesis, Ibn Zohr University, FLHS, Agadir
- Idrissi A (2000) On Berber plurals. In: Lecarme J, Lowenstamm J, Shlonsky U (eds) *Research in Afroasiatic grammar*. John Benjamins, Amsterdam, pp 101–124
- Idrissi A (2001) *Towards a root-and-template approach to shape-invariant morphology*. PhD dissertation, UQAM
- Inkelas S (1994) *The consequences of optimization for underspecification*. Ms. University of California, Berkeley. [Rutgers optimality archive#40]
- Itô J, Mester A (2003) Japanese morphophonemics- Markedness and word structure. The MIT Press, Cambridge, MA
- Itô J, Mester A, Padgett J (1995) Licensing and underspecification in optimality theory. *Ling Inq* 26: 571–613
- Jebbour A (1985) *La Labiovélarisation en berbère- Dialecte tachelhit- (Parler de Tiznit)*. C.E.U.S memoir, Mohammed V University, Faculty of Letters, Rabat
- Jebbour A (1988) *Processus de formation du pluriel nominal en tamazight. (Tachelhit de Tiznit)- Approche non-concaténative*. D.E.S. thesis, Mohammed V University, Faculty of Letters, Rabat
- Jebbour A (1991) *Structure morphologique du nom et problème de la voyelle initiale des noms en tachelhit, parler de Tiznit (Maroc)*. *Etudes Doc Berbères* 8:27–51
- Jebbour A (1996) *Morphologie et contraintes prosodiques en berbère (Tachelhit de Tiznit)- Analyse linguistique et traitement automatique*. Doctorat d’Etat thesis, Mohammed V University, Faculty of Letters, Rabat
- Kager R (1999) *Optimality theory*. Cambridge University Press, Cambridge
- Kawahara S, Sano S (2014) Identity avoidance and rendaku. *Proceedings of phonology 2013*. http://ling.auf.net/lingbuzz/001947/current.pdf?_s=rSnoRzt9b_ujoxo
- Kisseberth CW (1970) On the functional unity of phonological rules. *Ling Inq* 1:291–306
- Kisseberth CW (2011) Conspiracies. In: van Oostendorp M, Ewen CJ, Hume EV, Rice K (eds) *The Blackwell companion to phonology*, vol 1. Wiley-Blackwell, Malden/Oxford, UK, pp 1644–1665

- Krämer M (2003) Vowel harmony and correspondence theory. Walter de Gruyter, Berlin
- Kurisu K (2001) The phonology of morpheme realization. PhD dissertation, University of California, Santa Cruz
- Ladefoged P, Maddieson I (1996) The sounds of the world's languages. Blackwell, Oxford
- Lahrouchi M (2013) Templates, markers and syntactic structure in Tashlhiyt Berber. *Lingua* 133: 53–72
- Lasri A (1991) Aspects de la phonologie non-linéaire du parler berbère chleuh de Tidli. Doctoral dissertation, Université de la Sorbonne Nouvelle, Paris III
- Lombardi L (1998) Evidence for Max-feature constraints from Japanese. Ms. University of Maryland, College Park. [Rutgers optimality archive#247]
- McCarthy J (1986) OCP effects: gemination and antigemination. *Ling Inq* 17:207–263
- McCarthy J (2002) A thematic guide to optimality theory. Cambridge University Press, Cambridge
- McCarthy J (2007) What is optimality theory? *Lang Ling Compass* 1(4):260–291
- McCarthy J (2008) Doing optimality theory. Blackwell, Malden
- McCarthy J, Prince A (1993) Prosodic morphology I: constraint interaction and satisfaction. Ms. University of Massachusetts, Amherst and Rutgers University
- McCarthy J, Prince A (1994) The emergence of the unmarked: optimality in prosodic morphology. *NELS* 24:333–379
- McCarthy J, Prince A (1995) Faithfulness and reduplicative identity. University of Massachusetts occasional papers in linguistics 18: papers in optimality theory. [Rutgers optimality archive#60]
- McCarthy J, Prince A (1999) Faithfulness and identity in prosodic morphology. In: Kager R, van der Hulst H, Zonneveld W (eds) The prosody-morphology interface. Cambridge University Press, Cambridge, pp 218–309
- McCarthy J, Prince A (2004) The emergence of the unmarked. In: McCarthy JJ (ed) Optimality theory in phonology- a reader. Blackwell, Malden, pp 483–494
- Moktadir K (1989) The passive form in Tashlhiyt Berber- a prosodic approach. D.E.S. thesis, Mohammed V University, Faculty of Letters, Rabat
- Moran S, McCloy D (eds) (2019) PHOIBLE 2.0. Jena: Max Planck Institute for the Science of Human History. Available online at <http://phoible.org>. Accessed 15 Feb 2023
- Morén-Duolljá B (2011) Vowel place. In: van Oostendorp M, Ewen CJ, Hume EV, Rice K (eds) The Blackwell companion to phonology, vol 1. Wiley-Blackwell, Malden/Oxford, UK, pp 440–464
- Nevins A (2010) Locality in vowel harmony. The MIT Press, Cambridge, MA
- Ní Chiosáin M, Padgett J (2001) Markedness, segment realization, and locality in spreading. In: Lombardi L (ed) Segmental phonology in optimality theory: constraints and representations. Cambridge University Press, Cambridge, pp 118–156
- Odden D (1994) Adjacency parameters in phonology. *Language* 70:289–330
- Padgett J (2002) Feature classes in phonology. *Language* 78:81–110
- Padgett J (2011) Consonant-vowel place feature interactions. In: van Oostendorp M, Ewen CJ, Hume EV, Rice K (eds) The Blackwell companion to phonology, vol 3. Wiley-Blackwell, Malden/Oxford, UK, pp 1761–1786
- Prince A, Smolensky P (1993/2004) Optimality theory: constraint interaction in generative grammar. Blackwell, Malden/Oxford, UK. [Revision of 1993 technical report, Rutgers University Centre for Cognitive Science. ROA#537]
- Rose R (2011) Long-distance assimilation of consonants. In: van Oostendorp M, Ewen CJ, Hume EV, Rice K (eds) The Blackwell companion to phonology, vol 3. Wiley-Blackwell, Malden/Oxford, UK, pp 1811–1836
- Rose S, Walker R (2011) Harmony systems. In: Goldsmith J, Riggle J, Yu ACL (eds) The handbook of phonological theory, 2nd edn. Wiley-Blackwell, Oxford, pp 240–290
- Rubin AD (2008) The subgrouping of the Semitic languages. *Lang Ling Compass* 2(1):79–102
- Saib J (1982) Initial vowel syncope and reduction in Tamazight-Berber nouns. *Lang Lit* 2:159–184
- Saib J (1986) Noun pluralization in Berber: a study of internal reconstruction. *Lang Lit* 5:109–133
- Selkirk E (1993) [Labial] relations. Ms. University of Massachusetts, Amherst

- Smolensky P (1993) Harmony, markedness, and phonological activity. Handout from *ROW-1*, revised. [Rutgers optimality archive#87]
- Smolensky P (1995) On the internal structure of the constraint component Con of UG. Handout of talk, UCLA, April 7, 1995. [Rutgers optimality archive#86]
- Smolensky P (1997) Constraint interaction in generative grammar II: Local conjunction (or, random rules in universal grammar). Talk presented at the Hopkins Optimality Theory Conference
- Steriade D (1987) Locality conditions and feature geometry. *NELS* 17(2):595–617
- Suzuki K (1998) A typological investigation of dissimilation. PhD dissertation, The University of Arizona
- van de Weijer J (2011) Secondary and double articulation. In: van Oostendorp M, Ewen CJ, Hume EV, Rice K (eds) *The Blackwell companion to phonology*, vol 1. Wiley-Blackwell, Malden/Oxford, UK, pp 694–710
- van der Hulst H, van de Weijer J (1995) Vowel harmony. In: Goldsmith J (ed) *The handbook of phonological theory*. Blackwell, Cambridge, MA, pp 495–534
- Walker R (2012) Vowel harmony in optimality theory. *Lang Ling Compass* 6(9):575–592
- Yip M (1995) Repetition and its avoidance: the case of Javanese. Ms. University of California, Irvine. [Rutgers optimality archive#83]
- Yip M (1998) Identity avoidance in phonology and morphology. In: Lapointe S, Brentari D, Farrell P (eds) *Morphology and its relation to phonology and syntax*. CSLI publications, Stanford, pp 216–246
- Zsiga EC (2011) Local assimilation. In: van Oostendorp M, Ewen CJ, Hume EV, Rice K (eds) *The Blackwell companion to phonology*, vol 3. Wiley-Blackwell, Malden/Oxford, UK, pp 1919–1942

Acquisition in a Multilingual Context: English Mid and Low Vowel Contrasts by Native Speakers of Tashlhit

3

Nancy C. Kula and Nabila Louriz

Contents

Introduction	56
Studies of L2 and L3 Phonology	57
Role of L1 and L2 Transfer in L3 Acquisition	59
Combined Transfer	60
The Multilingual Context: Tashlhit Speakers' Linguistic Repertoires	60
Tashlhit (L1)	61
Moroccan Arabic (MA)	62
Modern Standard Arabic (MSA)	62
French	63
English	63
The Study: Hypotheses and Predictions	64
Participants	65
Perceptual Experiments	66
Results	67
Production Experiments	69
Discussion	73
Role of L1 Transfer and Other Languages	73
Comparing Across Systems in Multilingual Learning	76
Implications for Language Teaching and Future Work	78
Conclusion	80
References	81

N. C. Kula (✉)

Leiden University Centre for Linguistics, Leiden, The Netherlands

University of the Free State, Bloemfontein, South Africa

e-mail: n.c.kula@hum.leidenuniv.nl

N. Louriz

University of Hassan II of Casablanca, Casablanca, Morocco

e-mail: nabila.louriz@univh2c.ma

Abstract

Tashlhit learners of English learn English in a multilingual context where Moroccan Arabic, Modern Standard Arabic, French as well as other Amazigh dialects are spoken. Perception and production experiments show that in the acquisition of mid and low vowels of English, contrasts are acoustically assimilated into composite vowels that differ both from the L1 and the other languages in the multilingual linguistic repertoires of speakers. This happens although the target mid vowels are present as allophones in MA and MSA and as phonemes in French. It is argued that the prevalence of three-vowel systems in this context has a confounding effect on the acquisition of new contrasts.

Introduction

Many African contexts are well known for their prevalent multilingualism where speakers have at their disposal rich multilingual repertoires that feed complex but common day-to-day linguistic interactions. The switch from one language to another in these contexts is both smooth and unconscious but also, dictated by the context. How language acquisition works in these situations is less well understood although studies of third language acquisition have been conducted and provide the closest approximation of what we might expect in a multilingual context.

There are a number of factors that make investigating the acquisition of further languages beyond the L1 and L2 interesting, dynamic, complex, and challenging at the same time. It involves more variables than L2 acquisition presents and magnifies the way the number of previously acquired languages affect the various possible interactions in the learner's linguistic knowledge. Researchers are also interested in whether the order of acquisition affects the acquisition configurations. In case of speakers of Arabic and French acquiring English, does it matter if French is L1 and Arabic is L2 or vice versa? Moreover, the typological distance between the known languages and the target raises more questions for investigation. Other factors like fluency, L2 status or other language status in the acquisition of additional languages, and frequency of use have all been shown to play a role in acquisition (De Angelis 2007; Falk and Bardel 2010, 2011; Rothman and Cabrelli Amaro 2010). Although the development of this area of research has been flourishing, acquisition of the phonology of additional languages remains an understudied area.

It has been argued that while L2 and L3 language learners might attain a native-like mastery in areas of grammar such as syntax and morphology, they almost never master target language pronunciation satisfactorily (Littlewood 1994). Thus, L2/L3 learners must construct their own phonological representations to attain the desirable pronunciation. Ard (1990) finds that the interlanguage representation is not necessarily the same as that of the native speaker of the target language. Louriz (2004, 2007) shows how the correct (and incorrect) placement of primary stress of Moroccan Arabic natives learning English as L3 is the result of a representation that is different from that of English. L3 learners rerank the constraints available to them in

their L1 and L2 to construct a new representation for English. [Louriz \(2004, 2007\)](#) couches her work within Optimality Theory and argues that highly ranked constraints in L1 and L2 are demoted in L3 grammar. For example, the constraint that requires the foot bearing primary stress to be aligned to the right edge of the prosodic word is demoted below Align-L (Ft', PwD), which is undominated in the learners' L1 and L2 to produce primary stress in English earlier in the word than it is in the languages they speak (Moroccan Arabic and French.)

L2 phonology literature argues that sounds that undergo perceptual assimilation are those in close proximity in acoustic space. For instance, studies show that i/e and u/o are deemed non-assimilable in L2 vowel acquisition ([Gordon 2007](#); [Ingram and Park 1997](#)) because they are at a far enough distance in acoustic space to not cause any interference. We investigate these same vowel pairings, in addition to others, with respect to Tashlhit learners of English.

This chapter will investigate the production and perception of English [e ɔ æ ʌ] by L1 Tashlhit learners of English as an additional language in a multilingual context. (We will use the term “additional language acquisition” to refer to the acquisition of languages beyond a second language and specifically acquisition in multilingual contexts. IPA is used for phonetic transcription/representation of vowels.) We will particularly examine the influence of L1 Tashlhit, as well as other languages in the linguistic repertoires of speakers, which in this case specifically includes Moroccan Arabic (MA), Standard Arabic (MSA), and French, on their acquisition of English mid and low vowel contrasts. We thus aim to trace the emergence of a vowel system in a multilingual learning environment.

The chapter proceeds as follows. Section “[Studies of L2 and L3 Phonology](#)” provides an overview of relevant studies in L2 acquisition and the so-called L3 acquisition; section “[The Multilingual Context: Tashlhit Speakers' Linguistic Repertoires](#)” discusses the multilingual linguistic context in which Tashlhit is spoken, examining the different languages used; section “[The Study: Hypotheses and Predictions](#)” introduces the study and presents the perception and production experiments and the results; section “[Discussion](#)” provides a discussion of the findings, seeking explanations and also considering broader implications for language teaching and future work; and section “[Conclusion](#)” provides some final conclusions.

Studies of L2 and L3 Phonology

[De Angelis \(2007: 11\)](#) proposes that “third or additional language acquisition refers to all languages beyond the L2 without giving preference to any particular language.” An L3 is thus a special case of the wider category of L2, and not necessarily language number three in order of acquisition ([Hammarberg 2001](#)). (What the term L3 continues to espouse though is the strict ordering of language acquisition and the assumed distinct grammars of these languages. We would like to explore more flexibility in the designation and roles of languages in multilingual contexts to capture more accurately the ease of transfer of use between the different languages that may not restrict these to totally distinct categories. In this sense “additional

language” is evoking a different categorization and conceptualization of languages than L3 does, even when it refers to the acquisition of more than three languages.)

Archibald (2021) compiles an overview in L2 phonology research pointing to the major findings in the field throughout its development. He summarizes the factors that have been explored in L2 phonology literature to explain acquisition as well as variation in production. These are L1 transfer (Trofimovich and Baker 2006), amount of experience (Bohn and Flege 1992), amount of L2 use (Guion et al. 2000), age of learning (Abrahamsson and Hyltenstam, 2009), orthography (Escudero and Wanrooij 2010; Bassetti et al. 2015), frequency (Davidson 2006), attention (Guion and Pederson 2007), training (Wang et al. 2003), as well as cross-language speech production (Flege 1995) and cross-language speech perception (Best and Tyler 2007), which are reported to affect acoustic and articulatory aspects of acquisition.

Other researchers relate the success in the acquisition of L2 to markedness (Parker 2012); marked structures are more difficult to acquire than unmarked ones. The seminal work of Eckman (1985) proposes that markedness differential is crucial rather than markedness continuum per se. An L2 structure that is marked and does not exist in the L1 would be difficult to acquire, whereas an unmarked structure that is absent in the L1 would be easier to acquire. Since unmarked structures are frequent in most of the world’s languages, Archibald (2021) raises the issue of how and when to relate structures that are easy to acquire to frequency or markedness. Various approaches to frequency have been introduced. A usage-based approach in Wulff and Ellis (2018) relates frequency to usage, whereas Wilson and Davidson (2009) relate it to frequency of input. As far as production is concerned, markedness has been reported to be more explanatory than frequency. On the other hand, it has also been reported that the most frequent structure is the most marked (Cardoso 2007).

Second and third language acquisition are now treated as distinct areas of research in the literature. While L2 learners have only one language to transfer from, L3 learners have more. This latter scenario of language learning involves more complexity in that it involves transfer from L1/L2 to L3 and vice versa. There are a few studies that have investigated phonetic or phonological transfer in L3. These have shown that having a wider range of phonetic and phonological knowledge, L3 learners possess a larger inventory of articulatory and perceptual awareness, which is supposed to make L3 learning easier. Tremblay (2010), for example, shows that multilingual learners do better in discriminating new sound contrasts. However, other studies contrast with this. Werker (1986) claims that early bilinguals and monolinguals are similar in perceiving new L3 contrasts, implying that at least early bilinguals show no advantage over monolinguals. Gallardo del Puerto (2007) observed Basque/Spanish bilinguals acquiring L3 English and found that bilingual proficiency is not predictive of L3 perceptual ability. There are thus mixed views on whether multiple languages offer advantage when acquiring additional languages.

Apart from general ease or difficulty of learning of an L3, a number of studies identify a range of factors that may affect L3 acquisition. These include: the actual or perceived typological distance between L2 and L3 (e.g., Rothman and Cabrelli Amaro 2010; Rothman 2010; Rothman 2011); L1/L2 status (e.g., Bardel and

Falk 2007; Falk and Bardel 2011; Hammarberg 2001); recency of acquisition and proficiency in L2 (e.g., Barkley 2010; Ringbom 2002; Williams and Hammarberg 1998); and psycho-affective factors (e.g., De Angelis and Selinker 2001). In the below discussion we consider studies particularly related to transfer in an L3 context.

Role of L1 and L2 Transfer in L3 Acquisition

There is lack of consensus regarding the role the L1 plays in L3 acquisition. Some researchers argue that L1 has more influence than L2 in the process of acquiring L3. Llisteri and Poch-Olivé (1987) conduct an acoustic analysis of L3 vowels produced by native speakers of Catalan and L2-Spanish. The results reveal that in that case L1 exclusively affects L3 production. In a study of L3 intonation, Ringbom (1987) finds that L2 transfer is rare especially since advanced L3 learners maintain their L1-based accent. Wrembel (2013) conducts a study where the participants' mother tongue was the dominant source of transfer. He found no trace of influence from L2. By contrast, Gut (2010) asserts that L1 plays no role in learning English rhythm and vowel reduction in L3 learners' production.

With respect to the role of L2, numerous studies suggest that it is mainly the L2 that frequently serves as a predominant source of transfer in L3 acquisition. Some research finds that it involves positive transfer in L3 acquisition, whereas others argue that it has a negative effect. Some studies (e.g., Williams and Hammarberg 1998) relate the degree of L2 transfer to psychotypology, recency of acquisition, and degree of proficiency.

Kamiyama (2007) finds that L2 orthography has a negative influence in both perception and production of L3 French vowels by L1 Japanese/L2 English speakers. Tremblay (2007) underlines positive L2 influence in voice onset time (VOT) of L1 English/L2 French learners of L3 Japanese. Llana et al. (2010) confirm these findings in another study of VOT of L1 English/L2 French learners of L3 Spanish.

Several explanations have been suggested for L2 transfer. Wrembel (2010) suggests that at the initial stage of third language speech production, L2 mechanisms are frequently reactivated blocking transfer from one's mother tongue (the L1). This is motivated by the resemblance in processes involved in non-native acquisition that L3 learners have already experienced. This might explain why Rivers (1979) shows L2 influence from a language that had nearly been forgotten. This is in line with the account made by Hammarberg and Hammarberg (2009) who relate the strong role of L2 at the initial stage of L3 acquisition to the learners' cognitive mode, which deactivates L1 transfer and relies on L2 to deal with foreign phonetic forms. They suggest that the role of L2 decreases at later stages of L3 acquisition. Supporting this, Wrembel (2010) found evidence for an increased L1 influence among L3 learners at advanced stages of acquisition.

Other studies relate L2 influence on L3 acquisition to psycho-affective factors (Edmondson 2001; Hammarberg and Hammarberg 2009). According to De

Angelis (2005) non-native languages are usually classified as “foreign language” category in learners’ minds thus simply creating a cognitive association between them. As one’s L1 does not sound “foreign” it is usually excluded and blocked from this association. De Angelis (2005) calls this process an “association of foreignness” (p. 42). It usually favors non-native transfer, giving L2 a privileged status. This phenomenon was observed quite early by Meisel (1983), who named it a foreign language effect.

Combined Transfer

Apart from transfer in L3 acquisition being either strictly from the L1 or the L2, Louriz (2004, 2007) introduces the term “dual transfer” to refer to cases where both L1 and L2 influence L3 in relation to the acquisition of stress. Similarly, De Angelis (2007) refers to combined transfer to express the same effect. A few later studies make similar observations. Amaro (2012) states that “the influence from multiple languages and universals is not seen in a single segment, but rather across production, whether simultaneously (e.g., Barkley 2010) or successively over the course of acquisition (e.g., Hammarberg and Hammarberg 2009).”

Benrabah (1991) studies the segmental acquisition of Algerian Arabic/French bilinguals acquiring L3 English and found that learners transfer the more complex system of each (i.e., Algerian Arabic consonants and French vowels) into L3 English. This implies that the system that provides a wider segmental choice/range is the learner’s reference in acquisition. Wrembel (2011) analyzed the VOT in the L3-French speech of L1-Polish and L2-English users showing that there was a combined cross-linguistic influence of both L1-Polish and L2-English on the participants’ L3 production in the study.

Also supporting combined transfer, Hammarberg (2001) introduces the Role Function Model to explain the combined influence L1 and L2 have in the course of L3 acquisition. On the one hand, L1’s fixed neuro-motor system controls the learner’s articulation during L3 acquisition. On the other hand, given L2-L3 foreign relatedness, L2 acts as a coping strategy that overrides L1 influence at the onset of acquisition. As L3 proficiency improves, there is less phonetic transfer from L2 and the influence of L1 comes to the surface. Hammarberg and Hammarberg (2009) support this proposal in their longitudinal L3 phonology study and find that while L3 learners start by transferring from L2, they transfer from both L1 and L2 as acquisition develops.

The Multilingual Context: Tashlhit Speakers’ Linguistic Repertoires

The central purpose of this chapter is to understand acquisition of an additional language in a multilingual context and shed light on whether we can identify what influences the phonological acquisition process in this context. Specifically, to find

out and explain the role that the languages involved play in the acquisition process. Our subjects are Tashlhit native speakers learning English in a context where they speak or interact with at least four other languages. We consider the linguistic landscape in Morocco and the context of acquisition of our learners, noting also that this is further enriched by varieties of the languages involved, particularly Moroccan Arabic and Amazigh varieties other than Tashlhit that we do not consider in this study. The multilingual language situation in Morocco affects all aspects of life in formal and informal sectors.

In this section we discuss the language background of the learners involved in this study and contextualize the status and usages of languages in Morocco, in general, and in the Amazigh speaking area, specifically. The Tashlhit native speakers in this study who are learning English, speak and are also exposed to Moroccan Arabic (MA), Modern Standard Arabic (MSA), and French. The role and uses of these languages in their linguistic repertoires are provided below.

Tashlhit (L1)

Tashlhit is a variety of the Amazigh language, which is a 5000-year-old Afro-asiatic (Hamito-Semitic) language (Sadiqi 2001). It has been used by Imazighen who were the first inhabitants of Morocco for many centuries. This language has survived throughout the successive civilizations and languages that arrived in Morocco (e.g., Latin, Phoenician, Arabic). Despite the lack of an exact census, most sources report that at least 60% of Moroccans are ethnically Imazighen and acquire Amazigh as a native language (Sadiqi 2001: 65). Tifinagh-IRCAM is the widely used script adopted in the early 2000s.

There are three dominant varieties of Amazigh spoken in Morocco. Tarifit is used in the Rif Mountains and the neighboring valleys, situated in the north of Morocco. Tamazight and Tashlhit are spoken in the High Atlas Mountains with Tamazight also spoken in the Middle Atlas and Tashlhit also spoken in the Anti-Atlas.

Like most Amazigh languages, Tashlhit has three-vowel phonemes {a, i, u}, which may be raised and retracted when tautosyllabic with an emphatic consonant resulting in the allophones {ɑ, e, o}. Given the basic inventory of the vowel system, the three vowels have a wide range in the vowel space from [a] to [æ]; [i] to [e], and [u] to [o] (Kossmann and Stroomer 1997). As far as schwa is concerned, its status is somewhat unclear and disputed with, e.g., Ridouane (2008, 2011) arguing for consonant syllabicity rather than epenthesis, while Bensoukas (2006, 2012) and Bensoukas and Boudlal (2012) adopt the view that schwa occurs in prosodic morphology. This differs from other Amazigh varieties – Tamazight and Tarifit – which have schwa (Dell and Elmedlaoui 1988; Coleman 1999; among others). (Syllable structure and schwa syllabification in Tashlhit are part of ongoing interesting debate that has produced different competing perspectives. See Ridouane (2016) for an excellent overview of Tashlhit phonology.) The absence/status of schwa in Tashlhit has no direct bearing on the current study, which is focused on full vowel contrasts.

Moroccan Arabic (MA)

Moroccan Arabic is the product of a contact situation between Amazigh and Eastern varieties of Arabic. While the former is the native language of the land, the latter was introduced with the Islamization of Morocco in the eighth century and, later on, with several waves of immigration from Eastern populations. It refers to the variety of Arabic used as lingua franca by most Amazigh Arab natives in Morocco. It is important to note that MA is not one homogeneous language but is composed of a number of dialects; Heath (2002) suggests at least 15 dialects.

In this respect, MA speakers can be divided into two categories. There are “old speakers,” who base their language practice on their local dialectal variety, and “new speakers” lacking this background, either due to an Amazigh home language or growing up in one of the urban “melting pots” without a particular dialectal profile (Maas 2011). This distinction is important in our study, because our learners are considered new speakers of MA who mainly only speak MA outside their homes. Most such speakers are anecdotally said to speak MA with their native Amazigh intonation.

As far as the sound system is concerned, Amazigh phonology is believed to have a strong influence on MA (Sadiqi 2001). This is most noticeable in the vowel system and the syllable structure that makes Moroccan Arabic syllables more complex than other Arabic dialects. The rural MA varieties naturally contain fewer Amazigh elements than the urban varieties, although some borrowed vocabulary is present in their lexicon.

It is generally accepted that MA consists of three-vowel phonemes {i, a, u} in addition to the central epenthetic vowel {ə}. However, there is a debate on quantity. Arabists tend to retain the difference between long and short vowels of Old Arabic, maintained in Modern Standard Arabic (a second language to literate MA natives), treating the central vowel [ə] as a short vowel in MA. Some scholars consider it to be epenthetic since it splits non-admissible consonant sequences (Benhallam 1990; Boudlal 2001; Maas 2011; Lahrouchi 2018). In addition, its phonetic description is also fuzzy; it is seen as schwa (Harris 1942; Benhallam 1980, 1990; Keegan 1986; Heath 1987; Boudlal 2001) or as a vocoid (Dell and El Medlaoui 2002) or even that its phonetic signal is to be interpreted as a mere audible consonantal release as shown by its phonetic signal (Gafos 2002).

Modern Standard Arabic (MSA)

Standard Arabic is generally used in education, media, and administration as well as in political and scientific discourse. It has the prestige of a written language in the sense that it is codified and standardized. However, it is not the mother tongue of any speaker. It is related to the Arabization policy in Morocco, whose objective, among others, is to develop and modernize it. The major function of MSA is to express the aspects of modern culture, which had been expressed only in French previously, and which now have the role of introducing the occidental culture into the country. For

this reason, MSA has concentrated on lexical innovation and translation of French terms (Ennaji 2001; Gago Gomez 2019 and references therein).

Unlike Classical Arabic, MSA is characterized by a new enriched lexicon and by borrowings from French, as well as English and Spanish. It is also distinguished by syntactic and stylistic changes and a phonology that is highly influenced by dialectal Arabic. Despite this, it counts as part of the linguistic knowledge of our learners that our participants would have learnt as part of schooling. MSA, like Tashlhit and MA, has three phonemic vowels {a, i, u} and their long counterparts {a:, i:, u:}. For our Tashlhit learners of English this offers no additional contrast in height apart from high and low. Like MA and Tashlhit, the six vowels are lowered and retracted in emphatic environments.

French

French has been strongly present in Morocco since the 1912 treaty, which made French the official language of Morocco. It was imposed as “the only language of civilization and advancement” (Bourhis 1982). It has since become the main language of education; with its position strongly strengthened by the school system, taught from primary school (p. 14).

French is used in government, business, and media alongside Arabic varieties. However, Moroccan speakers’ fluency in French varies according to learners’ level of education. As far as our study is concerned, French introduces to our learners a wider range of vowels than MA and MSA would have previously done. French has 12 oral vowels and four nasal vowels (Armstrong 2021) as shown below.

(1) French vowel system

i	y		u		
e	œ		o	ø	õ
ɛ	œ̃	ə	ɔ	ẽ	ã
a			ɑ		

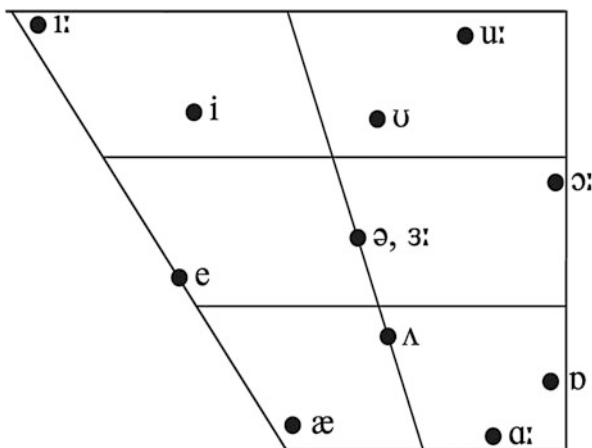
Thus, our learners have through French, four vowel heights contrasting the front and back distinctions as well as the dimension on rounded front vowels, in addition to a central vowel – schwa. There is further complexity in the system with also four phonemic nasal vowels, at least three of which are mid vowels.

English

In a globalized world, English is a popular language among the young generation with one of the most popular courses at university being English studies. English is taught in Morocco, although French is the national language. In state schools, students are introduced to English from the last year of secondary school (rather than of high school, in the last decade). We observe though that the day-to-day use of

English is still limited in the wider community with most interactions in English mainly in the context of learning the language. At least for the context of this study, learners were broadly mainly learning Southern Standard British English (SSBE) as the target taught in schools and university although they are also exposed to other dialects, particularly American English from the movie industry. Below is a vowel chart for SSBE based on Weckwerth (2021), which is derived from Jones (1918).

(2) Southern Standard British English Vowels (Weckwerth 2021)



Our study will specifically focus on the acquisition of the mid vowel contrasts {e, ɔ} and the low vowels {ʌ, æ} for which it has been argued, as noted above, that previous studies show no merger between i/e and u/ɔ. (Using standard vowel references as used in most phonetic and socio-phonetic work we are interested in the DRESS and LOT vowels for the mid vowels and the STRUT and the TRAP vowels. In most varieties of SSBE the STRUT vowel is a raised form of the START vowel and the distinction between these varies for different speakers (see Lindsey 2012 for detailed discussion). We treat the STRUT vowel as low although it is higher than the START vowel.) In the multilingual context of acquisition we investigate there are competing systems between MA and MSA, on the one hand, and French, on the other, that can have some interesting interaction with acquisition that we will explore in more detail below.

The Study: Hypotheses and Predictions

The goal of the study is to investigate the acquisition of the English mid vowels—front and back {e, ɔ} and the low vowels {ʌ, æ}. Given the number of vowels in the range of languages spoken by the participants, these vowels should be reasonably easy to acquire based on the following hypotheses.

- (i) There will be little interference from the three-vowel systems in Tashlhit, MA, and MSA because the three vowels are in the corners of the vowel space and far

enough apart to not result in interference with the English mid vowels, particularly {e and ɔ}.

- (ii) On the basis of French vowels, we expect learners to learn the English vowels easily, but perhaps have some interference from French and for the English vowels to be close to the French vowels in acoustic space, because the French vowels are exemplars of mid vowels in the learners' vowel space.
- (iii) The positions of the English vowels for the Tashlhit learners of English will be different from those of native speakers because of the multilingual repertoires they hold.

Below are the vowels relevant to the study and indications of which of these are present in the languages that the participants use. (A tick indicates the presence of a vowel category/type in a language while a cross indicates its absence. A tick in parenthesis shows that there is some allophonic variation for these vowels so that they are contextually present, primarily as a result of vowel lowering in the context of emphatics. Please note that while the English vowel categorization and descriptive labeling is used in the top row, we do not mean to suggest that where vowels are present in two languages, e.g., [i] in MA and English, that these are phonetically identical, hence our use of vowel "category/type" here. This hurdle is overcome in the detailed discussion to follow when the experimental results are considered. For [i:] and [o:] these are short where they occur in Tashlhit, MA, and MSA.) The target language English is the last one in the list (Table 3.1).

Much work in L2 acquisition has shown that accurate perception precedes and entails accurate production (Flege 1995, among others). Based on hypothesis (ii) we predict that Tashlhit learners of English will have no difficulty in the perception and production of the target English vowels because of their experience with peripheral mid vowel contrasts in French. We therefore conducted perception and production experiments to test the hypotheses given above. The details of the study and findings are given below.

Participants

The participants of the study were all students at the department of English language and literature at the University of Ibn Zohr in Agadir. They consisted of 20 native speakers of Tashlhit equally divided between males and females. Their age range

Table 3.1 Vowels in the multilingual context of Tashlhit learners of English compared to English

	i:/i FLEECE	ɪ KIT	e DRESS	ɑ:/ɑ START	ʌ a	æ TRAP	ʌ STRUT	o:/o NORTH	ɔ LOT	u:/u: FOOT
Tashlhit	✓	x	(✓)	x	✓	(✓)	x	(✓)	x	✓
MA	✓	x	(✓)	x	✓	(✓)	x	(✓)	x	✓
MSA	✓	x	(✓)	x	✓	(✓)	x	(✓)	x	✓
French	✓	x	✓	✓	✓	x	x	✓	✓	✓
English	✓	✓	✓	✓	x	✓	✓	✓	✓	✓

was between 19 and 23 years (mean 21). They all originate from the Southern Tashlhit speaking regions of Morocco, namely, Agadir, Chtouka Ait Baha, Tata, and Tiznit. At the time of testing, they were living in Agadir, and none reported any hearing deficits.

All participants had started learning English in the final year of secondary school and were still learning English at university. They grew up with Tashlhit as their only mother tongue spoken at home. At the age of 6–7 years old, they started learning MA. This is about school age, when they might have teachers and/or peers who are not Tashlhit speakers. They speak MA very fluently, like natives. Around the same age they were exposed to Modern Standard Arabic in the classroom. Two years later, they started learning French in primary school. They are advanced speakers of French and are able to use the language with ease although we did not test their level. The information regarding when they started learning/using a language are based on self-reporting, but they also follow patterns, which are consistent with the introduction of these languages in school. At the time of testing, they had been learning English for about 4–5 years (with an average of 4 h a week in the first 3 years). Based on the duration of learning and their proficiency we consider their English to be at about lower intermediate level. We conducted both perception and production tasks that lasted about 15 min each. Participants were compensated with a small token financial incentive.

Perceptual Experiments

Two experiments were conducted to test whether Tashlhit natives also speaking MA, MSA, and French can perceive selected mid and low vowel contrasts in English. The experiments consisted of discrimination and identification tasks.

The discrimination task consisted of an AX experiment that had 200 tokens with the five target vowel pairs given in a word {i/e, u/o, e/æ, ɪ/æ, ʌ/æ}. A native British English female speaker was recorded producing the items. Two repetitions of each item were retained as experimental stimuli. Instructions were given orally in Moroccan Arabic as to what each task entailed. Participants were tested individually in a quiet room. They received the speech stimuli through professional quality covering headphones. On each AX trial, participants were presented with two stimuli and had to indicate whether they were the same (AA) or different (AB). They indicated this by writing their response (S or D) on an answer sheet. The inter-stimulus interval was set to 3 s, and the inter-trial was 8 s.

For the identification task, informants were required to fill in blanks in C_C contexts, e.g., p_t and t_k based on their perception of nonce word CVC stimuli. Since there are fewer vowel symbols than there are vowel sounds, participants were asked to complement their response in the identification with an example of a word that contained the same vowel as they had identified, i.e., that rhymed. This was also a way of ensuring that they paid particular attention to the vowel. The vowels identified by the participants were transcribed. There were six target vowels in total with two repetitions. Two hundred and forty tokens were collected. Reaction

times were not measured in both tasks, but participants did not spend longer than 20 s on any one item. Participants were allowed to ask for any clarification if they needed it. This was taken up in very few cases.

Results

Results are displayed in Figs. 3.1 and 3.2. The striking observation is that subjects performed better in the discrimination task than in the identification task.

Figure 3.1 shows the overall discrimination accuracy for the five pairs of vowels. The learners performed well in discriminating between the vowel pairs {i/ε, u/ɔ, ε/æ, ɪ/æ, ʌ/æ}. Their correctness rate was near ceiling (95%) approaching that of native controls. However, they encountered difficulty discriminating between ʌ/æ (48%).

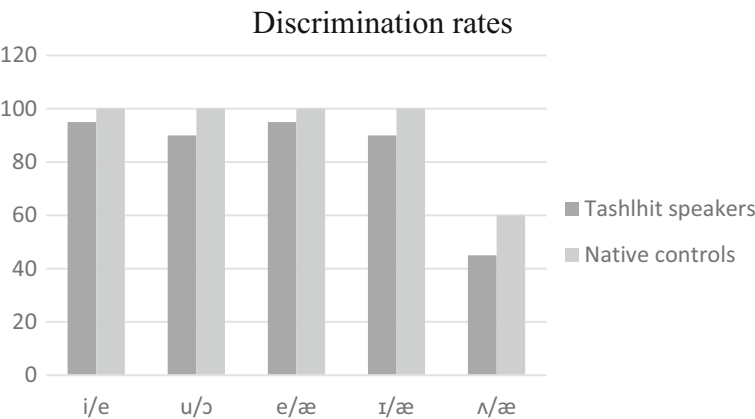


Fig. 3.1 Correct discrimination rates per vowel pair

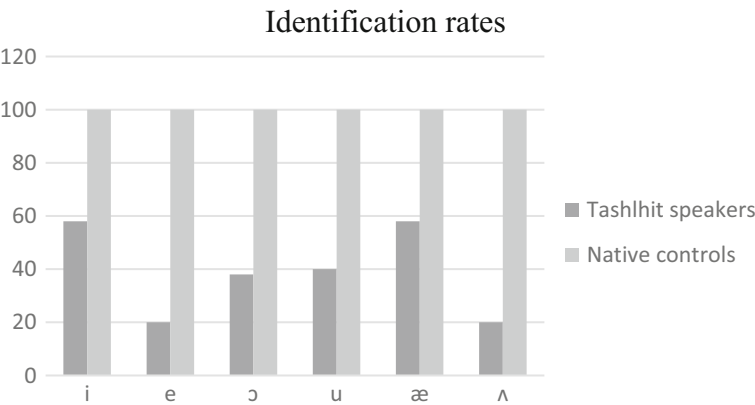


Fig. 3.2 Correct identification rates per stimulus vowel

The two groups of vowels are distinct in one major characteristic. The first four pairs differ in F1 only, whereas the last one differs in both F1 and F2. This suggests that the learners are more capable of discriminating vowels differing in F1 value only but have more difficulty with those where both the F1 and F2 values change. This results in the specific difficulty of discriminating between /ʌ/ and /æ/. However, note that native controls also performed less well in the last pair of vowels (60%). Both groups had a decrease of 40% correctness rate in this pair, which has been argued to be a particularly subtle contrast even for native speakers as the data show (see Lindsey 2012, for further discussion).

As far as the identification task is concerned, Fig. 3.2 demonstrates the percentages of correct identification per stimulus vowel. All native controls reached the 100% ceiling of correctness. Tashlhit English learners were dramatically unsuccessful in identifying all vowels; their most correct performance was just above chance level. The immediate observation is that learners performed better in what they perceive as the corner vowels, i.e., in {i, u, æ} than in mid vowels and [ʌ], although the performance in [ɔ] is better than [e]. Within the first group, they performed the best in identifying /i/ and /æ/ (front vowels) with 58% accuracy than in /u/ (back vowel) where they had only 40% accuracy rate. In contrast, as far as mid vowels are concerned, subjects identified [ɔ] with less difficulty (40%) than [e] and [ʌ] (20%). Despite the overall poor performance, results reveal that Tashlhit English learners performed better in identifying the three corner vowels than mid or other low vowels.

We conclude that the Tashlhit English learners perceive the target contrasts (discrimination test) but their perception did not allow them to accurately identify the vowels in the contrast (identification test). We summarize these results in the following confusion matrix, which demonstrates the rate or percentage of confusion in both tasks between any pair of vowels of the six target vowels (Table 3.2).

The matrix shows poor identification of {e, ɔ, ʌ} (light gray cells) for which an error analysis shows that they are assimilated to {i, u, æ}, which are identified more correctly overall. Recall that {e, o, a} are allophones in both MA and MSA occurring in the vicinity of tautosyllabic emphatic consonants, whereas they are phonemes in French (Armstrong 2021). When the Tashlhit learners misidentify mid vowels and low [ʌ], this is an indication that they confuse them with other vowels implying that Tashlhit learners perceive vowels like /i/ and /e/ as exemplars of the same category.

Table 3.2 Confusion matrix

	i	e	ɔ	u	æ	ʌ
i	58	33	0	3	3	0
e	32	15	7	3	28	13
ɔ	0	0	40	38	8	8
u	0	0	30	43	5	3
æ	0	8	5	0	58	23
ʌ	10	0	15	5	45	20

Our findings are in this sense in line with results in previous works such as Pineda Mora et al. (2019) but contrast with, for example, Ingram and Park (1997) who show *i/e* as being non-assimilable in Korean and Japanese learners of English.

We consider the production experiments below. Given the perception results above, where learners face difficulty, we can expect that production will also not be optimal with {*e*, *ɔ*, *ʌ*} not likely to be significantly different from {*i*, *u*, *æ*} in the Tashlhit English learners' productions.

Production Experiments

Production experiments were composed of three tasks consisting of target mid and low vowels. The first one was production of target words in a list presented in a carrier phrase produced at their own pace. Subjects were asked to produce the six target vowels in C_C nonce words in a carrier phrase. Two bilabial stops are used in the onset position {*b*, *p*} and three consonants at different manners in the coda position: {*m*, *t*, *s*}. None of these are close to emphatics that have a lowering effect. Thus *b_m*, *p_t*, and *b_s* created non-words totaling 260 tokens. In the second task, subjects were presented with a print-out of the experimental stimuli consisting of an extract from the "rainbow passage" (Fairbanks 1960). Participants read the passage out loud at their own pace. The last task was a CVC word-chain game, where participants had to create a string of words by adding a word progressing from one participant to the next. We report results from the first two tasks. (We do not analyze the data from this final task, which was intended to provide freer non-scripted productions as there were a number of pauses with participants also mainly unable to recall the preceding words. The task, however, worked very well as an ice breaker before the two production tasks that we report on here.) The experiments were recorded using zoom H4 with a sampling rate of 44 Hz.

Data were phonetically analyzed using Praat (Boersma and Weenink 1992–2022) looking at the formant values of the vowels produced and also using Praat to plot the vowel ranges. Measurements of the first and second formant frequencies (F1 and F2) were extracted at vowel midpoint using an automated procedure in Praat. To reduce individual differences, the Bark Difference Metric normalization was applied to F1 and F2 and was used in the figures presenting the vowel productions.

Results

Figure 3.3 shows the position of the English vowels {*i*, *e*, *u*, *ɔ*, *ʌ*, *æ*} as produced by multilingual Tashlhit learners of English. Results show a merger between {*i/e*} and {*u/ɔ*} although the standard deviation of the mid vowels {*e*, *ɔ*} are slightly less than and contained within those of {*i*, *u*}, respectively. Contra to expectation, [æ] and [ʌ] emerge as distinct.

Recall, as is well established, that the lower the F1 value, the closer the tongue is to the roof of the mouth, i.e., the higher the vowel, and the higher the F2, the more fronted the vowel. This is proportional to the frontness/backness of the highest part of the tongue. Examining the production range of [i] and [e], Fig. 3.4 demonstrates

vowel space is unquestionable. We see this in the separated picture in Fig. 3.4. There is similarly merger also between {u, ɔ} with [ɔ] contained within [u]. We discuss these results further in the extracted picture in Fig. 3.5.

Figure 3.5 demonstrates that {u, ɔ} merge, i.e., their production ranges overlap to a great extent implying confusion on the part of Tashlhit speakers. What is striking in contrast to i/e is that the range of u/ɔ is significantly larger with F2 ranging from 550 Hz to 2300 Hz and F1 from 300 Hz to 610 Hz. This indicates wide variation and range of vowels in production, although note that the concentration is mainly between 1000 Hz and 1600 Hz on F2.

For the ʌ/æ pair, Fig. 3.6 makes it evident that these vowels are mainly treated as distinct although there is an area of partial overlap. Their F2 ranges between 950 Hz and 2250 Hz with an overlap between 1400 Hz and 2000 Hz, and F1 from 300 Hz to 700 Hz with an overlap between 650 Hz and 450 Hz. This shows that production varies and merger is only partial.

In order to further analyze and better understand the observed variation, we divided the production data further into male and female speakers, which reveals contrasting patterns of performance as shown in Figs. 3.7 and 3.8.

Firstly, male production occupies a wider F2 (and to a less degree F1) range in contrast to female speakers. Secondly, the production range of peripheral and mid vowels in males is lowered and retracted compared to female production. And finally, unlike in female production, almost all vowels are partly or totally merged in male pronunciation, i.e., while there is vowel overlap for both males and females, there is less distinction between the vowels in male compared to female speakers.

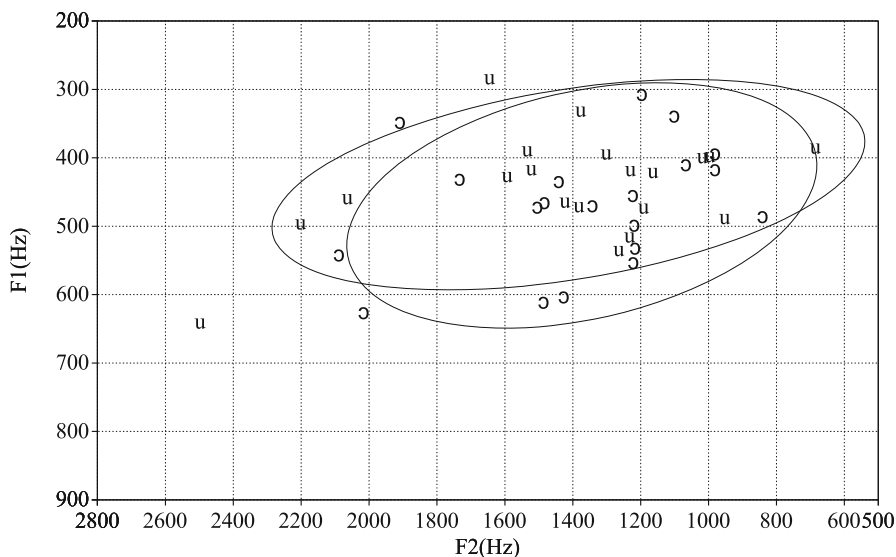


Fig. 3.5 u/ɔ production range

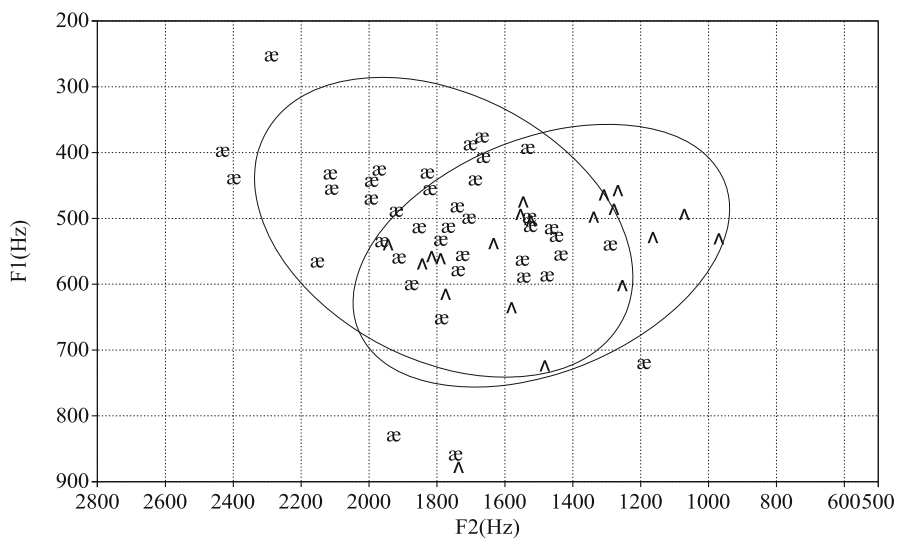


Fig. 3.6 /Λ/æ production range

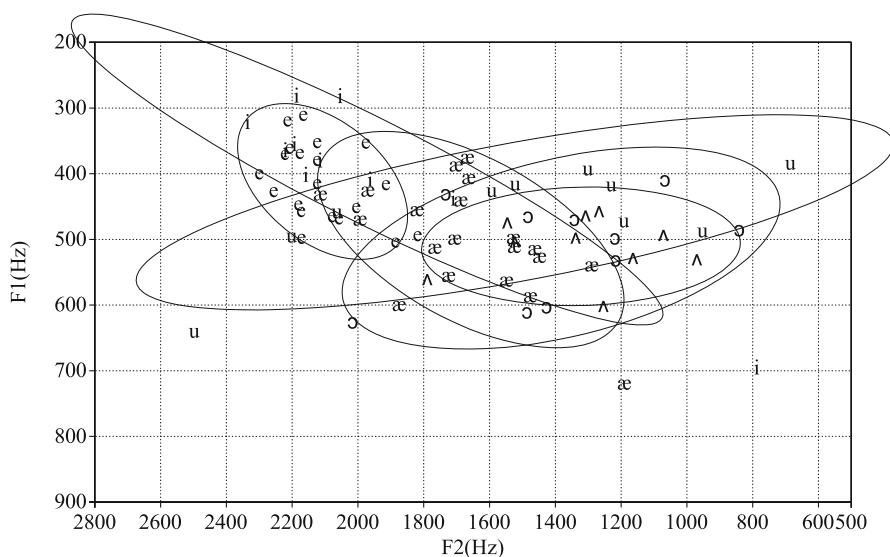


Fig. 3.7 Male Tashlhit English vowels

The production range of [i] in female speech is much narrower (higher and fronter) than in male pronunciation. The F2 value in females varies from 2100 Hz to 2250 Hz and F1 from 370 Hz to 320 Hz, whereas in males it is between 1700 Hz and 2500 Hz for F2, and 600 Hz to 180 Hz for F1. [e] production space seems to be totally different in male and female productions. This vowel is much lower and retracted in

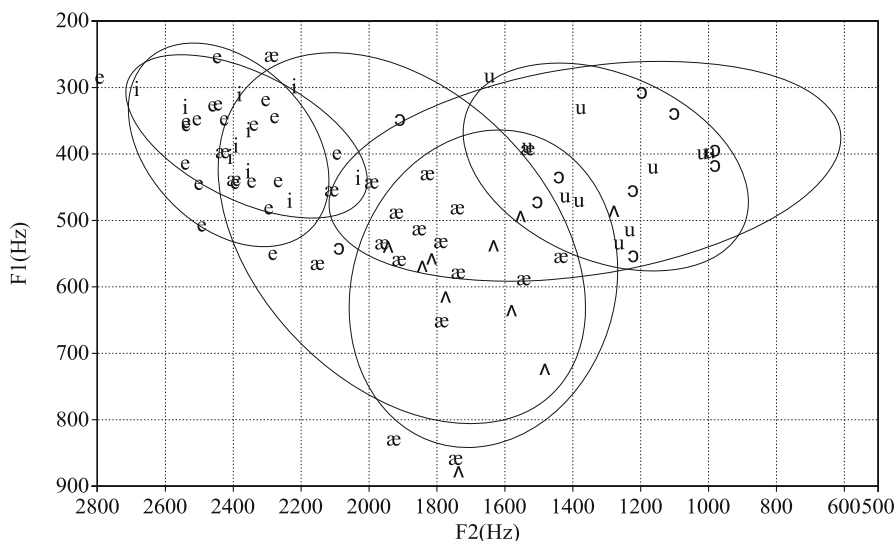


Fig. 3.8 Female Tashlhit English vowels

male speech. [u] in the female production range is also much smaller in terms of F2. By contrast, male production of [u] is spread across the vowel space (while there is not much difference in F1).

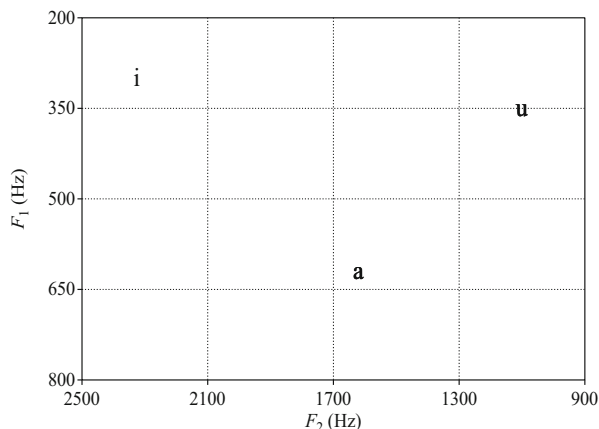
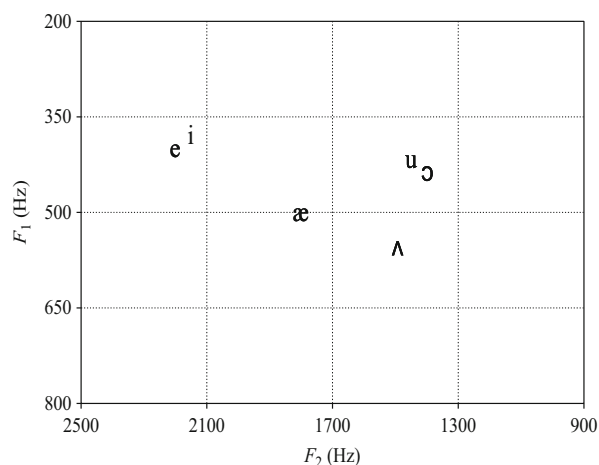
Female production range of [æ] is more concentrated (shorter F2) than in males. Thus overall, although there is an expected lower pitch in male than in female speakers, the difference in range is unexpected. Thus, at least in this sample of learners, female speakers do much better than males in moving toward distinguishing the vowels in the vowel space, although even in this case we see a merger between u/ɔ and i/e in female speakers but less so for æ/ʌ. There are no doubt important sociolinguistic factors that may have led to this division but which we do not pursue in the present paper.

Discussion

Role of L1 Transfer and Other Languages

In order to examine any influence of Tashlhit learners' mother tongue on their perception and production of English vowels, it is necessary to shed light on the acoustic values of the formant frequencies for L1 Tashlhit vowels (Ridouane 2014) and compare them to the mean value of their production of English vowels (Fig. 3.10). Compare Figs. 3.9 and 3.10 below. (These vowel plots are produced using Praat (Boersma and Weenink 1992–2022).)

Figure 3.9 shows that L1 vowel positions for Tashlhit are spaced low on F1 and wide apart on F2 for i/u and with high F1 for the low vowel /a/. Figure 3.10 shows

Fig. 3.9 L1 Tashlhit vowels**Fig. 3.10** Tashlhit English vowels

the position of the English vowels as produced by Tashlhit speakers, here showing the mid and low vowel targets as well as {i, u}. The comparison shows that Tashlhit speakers' English vowels [u] and [ʊ] are fronter and lower than the L1 [u] and the productions of [i] and [e] are lower and less front than the L1 [i]. [æ] is less low than L1 [a] but [ʌ] is distinct from [æ]. What would be six distinct vowels of English emerge as four vowels, in a pattern similar to the L1 but creating a smaller triangle of the merged L2 vowels plus [æ] that is contained within that of the L1 in acoustic space.

This demonstrates that the target mid vowels are misidentified with the target English vowels acoustically assimilated into composite vowels that differ both from the L1 /i/ and /u/, on the one hand, and from English i/e and u/ʊ, on the other.

Unexpectedly, French, one of the languages spoken by the participants, does not seem to figure at all in these productions in that it does not appear to aid productions to be closer to the target. Although it is tempting to view this as evidence of the supremacy of the L1, we would like to relate it to the multilingual context of acquisition. In this case, out of all the languages spoken and used, three of these – Tashlhit, MA and MSA – have three-vowel systems where mid vowels only occur as allophones. This has two effects. Firstly, the recurrence and dominance of three-vowel systems in the linguistic repertoires of these learners establishes these systems in a more compounded way that has a cumulatively stronger effect on the acquisition of further languages. Secondly, since mid vowels in these languages only occur as a result of a phonological context (presence of emphatics) that does not occur in English, and was therefore not present in the stimuli, the strong cues for mid vowel production are absent, and therefore perception is skewed toward merger as the data show.

Comparing F1 and F2 in Figs. 3.9 and 3.10, we observe that the greatest shift both from Tashlhit and English is seen in the F2 values. This is detailed in Table 3.3 below.

The F2 for [æ] in the learners’ English productions is greater than in both the L1 and the target language, although not very different from the L1. It is not clear at all why the learners make any change here since maintaining the L1 would be closer or indeed at the target but recall that there are more language formant values that speakers are exploiting and which must feed into their English productions. This means that looking at only the L1 does not provide the best insight and these findings thus support combined transfer (Louriz 2004, 2007; De Angelis 2007). In addition, there is also likely a foreignness effect where learners cognitively assume they need to make some change in the production of the newly being learnt language. But overall, in the absence of a detailed statistical analysis, learners are in a good range toward [æ]. For both productions of English [i] and [ɔ], the movements to a more intermediate position suggest some perception of the English F2 or more likely are the result of adjustments triggered by a fuller vowel space in the multilingual context. (Unfortunately, we did not measure these speakers own French productions to allow us to gauge how these relate to the English productions but make a comparison to native French productions based on formant values in the literature, see discussion in section “Comparing Across Systems in Multilingual Learning”)

On the other hand, the over increase in F1 for English [i, ɔ] results in vowels that are closer to mid vowels. We assume this is influenced by the assumption of a

Table 3.3 Comparison of F1/F2 between L1 Tashlhit and English productions

	F1/F2	Tashlhit L1	English	Tashlhit English
a	F1	620	667	513 (=æ)
	F2	1620	1565	1780
i	F1	300	296	390 (=e)
	F2	2325	2241	2186
u	F1	350	386	454 (=ɔ)
	F2	1100	1587	1392

smaller vowel system but also fed by perception and the expectation of new foreign vowels. This however, as the data show, does not affect F1 and F2 equally. Learners can manipulate or make changes to F2 with less difficulty/effort than to F1, although this is directly related to the fact that at least based on the L1 there is less change needed in the F1 category. Within the multilingual context, however, we expect that there is much more competition in the vowel space that results in shifts in vowels where we would predict them to just be based on the L1. This is what we consider to be the motivation for the shift fed by perception as reported.

This is possibly further supported by the orthographic links between French and English – they are both written with roman script; and they share several cognates that are spelt with similar graphemes, although given the non-commutability between spelling/grapheme and phonetic outputs, where, e.g., <a> can have multiple phonetic outputs in English (e.g., æ, ʌ, eɪ, ɑ, ə), this is likely to result in mis-mappings. In this case the lack of orthographic distinction between [æ] and [ʌ] may block the learners' production and may have affected the identification task results. This sort of metalinguistic knowledge has been shown to affect L2 acquisition (Bassetti et al. 2015; Hamann and Colombo 2017).

While the overall results could be argued to be in line with studies that show that the L1 has the greatest influence on additional language acquisition as Wrembel (2013) argues, two factors have to be considered in this context. Firstly, the L1 is not unique in this multilingual context, which is replete with three-vowel languages, and secondly, the results are also likely influenced by the acquisition stage of learners at the intermediate level. As previously noted, Hammarberg and Hammarberg (2009) suggest that speakers initially rely on the L2 (or other non-first languages) in the initial stages of learning but then begin to shift back to the L1 in later stages of development. What these results do not show is the observation made in Benrabah (1991) where learners transfer the more complex to the less complex system as was argued for Algerian Arabic speaking learners of L3 English who contrasted transfer patterns between consonants (from Algerian Arabic) and vowels (from French).

Comparing Across Systems in Multilingual Learning

We end the data discussion by comparing across the vowel systems in the multilingual space of the learners. A caveat on this is that this is a simulated scenario as while we map the actual merged vowel productions of the Tashlhit subjects, we rely on the literature for the positions of MA, English, and French vowels. In this sense this is a comparison of the learners' English productions against the assumed targets of the other languages as standardly produced by native speakers. An important piece that we leave to future work is a representation of the learners' actual vowel space with their productions of the relevant vowels to provide the actual multilingual space. We nevertheless think what we represent here is an important yardstick for this future work, which will allow us to also understand how the other languages may have shifted from native speaker productions.

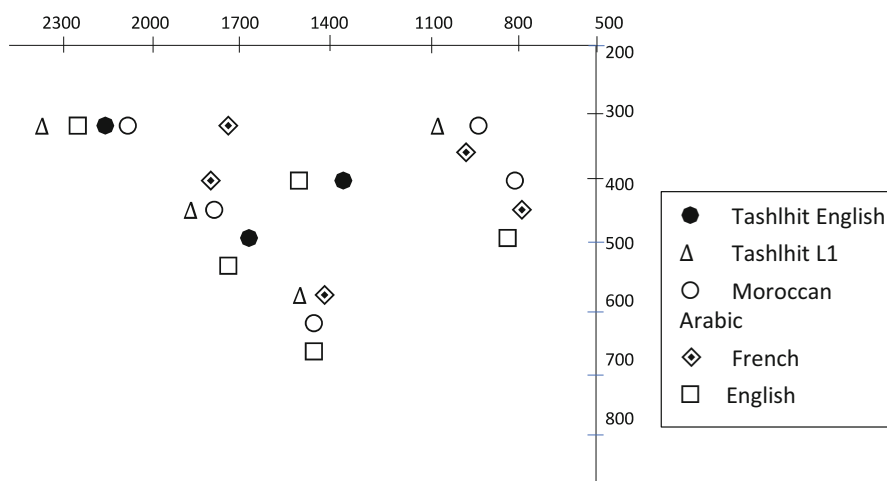


Fig. 3.11 A comparison of vowels in a simulated multilingual space

The mapped vowels in Fig. 3.11 below show English [ɪ e ə u ɔ] (Deterding 1997), French [a e i ɔ u] (Ouni and Laprie 2003), Moroccan Arabic [a i u] + [e o] allophones (Kenstowicz and Louriz 2009), Tashlhit [a i u] + [e] allophone, and the merged English productions of Tashlhit learners. The key to the vowel productions is given below.

Each filled circle point represents two vowels for merged English i/e and u/o and the single /æ/ for Tashlhit learners of English. At least compared to the position of native vowels, this illustrates that the vowels produced for English by the Tashlhit speakers are quite distant from the French vowels indicated by the diamond. This suggests that contrary to expectation based on a foreignness effect, French does not in fact influence the Tashlhit English vowels very much. Although the French vowels are close to at least one of the English vowels in each case, it is distant from another, therefore failing to make a contrast. In particular, the large change in F2 implies that the back vowel contrast is very poorly distinguished. Apart from u/a the acquired or emerging English vowels are close to MA and Tashlhit vowels to the extent that it seems that, at least for these vowel contrasts, a three-vowel system is what emerges. However, u/o using the mean value of tokens produced is much more central and not as back as all the other back vowels in all the other languages in the multilingual space. The data for English show significant /u/-fronting as widely attested in most recent work.

Models of L2 perception give supremacy to the L1 as the language through which the L2 is perceived (e.g., Boersma and Hamann 2009), and this then leads to mapping onto the relevant phonological forms. In the multilingual situation at hand, it is not so straightforward to decide which of the grammars in the linguistic repertoire are used for perception. If we consider that particular formant structures relate to particular vowels, and that these are the structures that the learner borrows from one or more of the languages in their repertoire, we end up with a quite

Table 3.4 Mixed formant signatures of Tashlhit learners of English

Vowel	F1	F2
e	= French, English	= MA
ɔ	= MA, French, English	= NONE
æ	= English	= MA, English

complex picture for Tashlhit learners of English. In this case, rather than complete F1/F2 pairs from one language being mapped to the English vowel output, the data seems to suggest perceptual assimilation of the composite vowels to different parts of vowels of different languages, i.e., formant signatures for a single vowel are potentially made up of formants from different vowels as depicted in Table 3.4.

These (developmental stage) results suggest at least that the Tashlhit speakers' perception was influenced by the multiple languages within their repertoire and not just the L1. In some cases, it's not even possible to uniquely identify which language the F1 or F2 formant is assimilated to and indeed it may be that there are multiple sources. For F2 and in particular for the back mid vowel /ɔ/ we see that this is an area of innovation where learners do not pattern with any of the languages in the repertoire.

Implications for Language Teaching and Future Work

The findings of this study have some implications or point to some considerations that can be made in English language teaching. Approaches to language teaching in non-native English-speaking contexts and in many parts of multilingual Africa, usually adopt an approach to language teaching where the English language classroom does not permit the use of other languages. Languages used outside of class are deemed to be inappropriate for the English classroom. This creates an artificial learning context in contrast to the way that language is used in the day-to-day lives of learners, with learners having to suppress the other languages in their repertoire when in class. More recent approaches to language learning (Eriling et al. 2017, 2021; McKinney 2020) challenge this approach and instead argue for multilingual language learning pedagogies where all the languages that learners have in their linguistic repertoire are treated as resources that enhance the learning of additional languages (Costley et al. 2023; Reilly et al. 2022, *In press*). In the results we have seen, it is surprising that despite the fact that the learners have French in their linguistic repertoire and French has at least two of the vowel contrasts they were trying to learn in English, with even a tense/lax distinction with {e, o, ε, ɔ}, speakers still had difficulty with mid vowels. Even though these vowels of French are not at the precise formant values for English, they would provide a better foundation and starting point for the learning of English rather than the L1, which seems to have the biggest impact on English vowel production and perception with little distinction between i/e and u/o.

A key finding in the present study is that the main cause of Tashlhit English learners' vowel merger in production is perceptual. Language teaching could

therefore provide training to help learners identify non-native contrasts since more accurate perception leads to more accurate comprehension and word recognition. The fact that the Tashlhit learners of English in this study did better in discrimination than in identification tasks shows that while they can distinguish minimal pairs, they cannot identify the vowels accurately and are therefore less able to use the correct vowels for their own productions, also as the results show. An important factor that also affects learners is variation in the input since they also listen to different varieties of English. Rather than be strict about variation in the input, we see this as a resource that teachers could use to teach pronunciation in class. As multilingual speakers, Tashlhit learners are adept at navigating different languages, and teachers using these skills that learners have, as well as their ability to contrast different languages, would allow learners to tap into their wider vowel inventory of French for the purpose of learning English. Although perfect native-like pronunciation is not the target and is indeed not expected in a multilingual context, teaching aspects of English pronunciation is useful to improve intelligibility particularly due to the non-transparent orthography of English, which can lead to mispronunciations as a result of analogies learners make based on orthography (e.g., the last syllable of the word “determine” is pronounced in analogy with “mine”). It would therefore be recommended that, in addition to teaching writing and reading, English pronunciation and sound structure are also included in both classrooms and in teacher training.

To further support the proposed multilingual approach to learning and teaching and to have a wider understanding of the linguistic competencies of the learners in their different languages future research in the following areas would be particularly beneficial. (i) A study where the production and perception of French vowels is investigated in this context would be significantly important as this would provide us with an understanding of whether the multilingual speakers have the resources available (here the knowledge of mid vowels at least) that they could then easily exploit in the learning of an additional language. Impressionistically, at least at the production level, the learners use of French exhibited a level of fluency that would support an advanced and stable representation of this language in their cognitive systems. (ii) To further test and verify the findings of this study, future work could replicate the study in this paper with L1 Moroccan Arabic speakers who are learners of English to consider whether another language with the same vowel system would yield the same results. Since Moroccan Arabic speakers are also usually speakers of French, any proposals on multilingual learning pedagogies would also be applicable in this context and beyond in the Maghreb. (iii) Future work could also importantly look at the role of allophones in the language learning context. As Table 3.1 shows, Tashlhit has variants of /e/ and /o/ that occur in restricted contexts, but these vowels are not phonologically contrastive in the language. Building on the early work of Eckman et al. (2003), this work could shed more light on the role of contractiveness and allophony in phonological acquisition, and the implication this has on the cognitive representation of allophones and also the acquisition of the phonological rules associated with such allophonic alternations in additional language acquisition. And finally (iv) while our findings provide the basis for further testable hypotheses in multilingual acquisition, either a longitudinal study or a staggered study of the

early, intermediate, and advanced stages of acquisition of English by multilingual Tashlhit speakers would provide further insight on the role that the different languages in the linguistic repertoires of speakers plays in the acquisition process and would allow us to better understand how learners navigate the different languages and evaluate what role each language takes at the different stages of acquisition.

Conclusion

We presented a case of acquisition of English in a multilingual language context of Tashlhit speakers who also speak Moroccan Arabic, Modern Standard Arabic, and French. We tested three hypotheses:

- (i) There will be little interference from the three-vowel systems in Tashlhit, MA, and MSA because the three vowels are in the corners of the vowel space and far enough apart to not result in interference with the English mid vowels, particularly {e and o}.
- (ii) On the basis of French vowels, we expect learners to learn the English vowels easily, but perhaps have some interference from French and for the English vowels to be close to the French vowels in acoustic space, because the French vowels are exemplars of mid vowels in their vowel space.
- (iii) The positions of the English vowels for the Tashlhit learners of English will be different from those of native speakers because of the multilingual repertoires they hold.

With hypothesis (i) we expected that there was likely to be little interference from the three-vowel systems in Tashlhit, MA, and MSA because the position of the three vowels in the vowel space could be deemed to be far enough away from the target vowels to result in any interference. To the contrary, we found that the Tashlhit speakers English mid vowels, particularly for {e, o}, merged these vowels with {i, u}, respectively. Based on hypothesis (ii) we expected that the learners could learn English vowels easily if they relied on French vowels, but at this intermediate stage of English acquisition the learners did not draw any benefits from French although their formant structures of the vowels produced do suggest some influence of French. In fact, supporting hypothesis (iii) it was shown that learners produced vowels that are influenced by all languages in their linguistic repertoire and their positions of their English vowels were different from those of native speakers because of the multilingual repertoires they hold. We think that our findings are unique to the multilingual learning context and that this provides support for a combined transfer approach to additional language acquisition where, because of the interaction between the different vowels used by the speakers, all the languages in the repertoire have an effect on the production and perception of any new language learnt.

Acknowledgments We express our sincere thanks to all the participants and especially Lasri Abdellah, Rachid Bouzougui, and Ahmed Naim from the University Ibn Zohr in Agadir for their help in recruiting the participants. Thanks to audiences at conferences and seminars in Bangor, Marrakech, and Poznan for discussion of different aspects of this work and to Silke Hamann for help with Praat scripting. We are also grateful to the reviewers of this chapter who have helped to sharpen the arguments made with some insightful suggestions. Funding for the fieldwork to conduct this research is gratefully acknowledged from the University of Essex DRIS funds from the Department of Language and Linguistics. All errors are our own.

References

- Abrahamsson N, Hylenstam K (2009) Age of onset and nativelikeness in a second language: listener perception versus linguistic scrutiny. *Lang Learn* 59(2):249–306. <https://doi.org/10.1111/j.1467-9922.2009.00507.x>
- Amaro CJ (2012) L3 phonology: an understudied domain
- Armstrong N (2021) Phonological variation and change in European French. Oxford Research Encyclopedia of Linguistics. <https://doi.org/10.1093/acrefore/9780199384655.013.491>
- Archibald J (2017) Transfer, contrastive analysis and interlanguage phonology. In: Thompson R, Kang O, Murphy J (eds) *The Routledge handbook of English pronunciation*. Routledge, London, pp 9–24
- Archibald J (2021) Ease and difficulty in L2 phonology: a mini-review. *Front Commun* 6:626529. <https://doi.org/10.3389/fcomm.2021.626529>
- Ard J (1990) A constructivist perspective on non-native phonology. In: Gass S, Schachter J (eds) *Linguistic perspectives on second language acquisition*. Cambridge University Press, Cambridge, pp 243–259
- Bardel C, Falk Y (2007) The role of the second language in third language acquisition: the case of Germanic syntax. *Second Lang Res* 23(4):459–484
- Barkley S (2010) The acquisition of orthographic-phonological correspondence rules in L2 and L3 Portuguese: error resolution, interference, and generalizability. PhD dissertation, University of Florida
- Bassetti B, Escudero P, Hayes-Harb R (2015) Second language phonology at the interface between acoustic and orthographic input. *Appl Psycholinguist* 36:1–6. <https://doi.org/10.1017/s0142716414000393>
- Benhallam A (1980) Syllable structure and rule types in Arabic. PhD Dissertation. University of Florida, Gainesville
- Benhallam A (1990) Native speaker intuitions about Moroccan Arabic stress. In: Jochen P (ed) *Maghreb linguistics*. Éditions Okad, Rabat, pp 91–109
- Benhallam A (1998) Syllable structure and rule types in Arabic. PhD dissertation, University of Florida, Gainesville
- Benrabah M (1991) Learning English segments with two languages. In: *Actes du XIIème Congrès International des Sciences Phonétiques*. Université de Provence, Service des Publications, Aix-en-Provence, pp 334–337
- Bensoukas K (2006) The emergence of the unmarked in Tashlhit round velar consonant dissimilation. In: Allati A (ed) *La Linguistique Amazighe: Les Nouveaux Horizons*. Publications of the Faculty of Letters, Tétouan, Tétouan, pp 76–118
- Bensoukas K (2012) The morphophonology of Tashlhit agentive nouns. *Lang Litt* 22:103–136
- Bensoukas K, Boudlal A (2012) The prosody of Moroccan Amazigh and Moroccan Arabic: similarities in the phonology of schwa. In: Borowsky O, Kawahara S, Sugahara M, Shinya T (eds) *Prosody matters: essays in honor of Lisa Selkirk*. Equinox, Sheffield, pp 3–42
- Best C, Tyler MD (2007) Nonnative and second-language speech perception: commonalities and complementarities. In: Munro MJ, Bohn O-S (eds) *Second language speech learning: the role of*

- language experience in speech perception and production. John Benjamins Publishing, Amsterdam, pp 13–34
- Boersma P, Hamann S (2009) Introduction. In: Boersma P, Hamann S (eds) *Phonology in perception*. Mouton de Gruyter, Amsterdam/New York, pp 1–19
- Boersma P, Weenink D (1992–2022) Praat: doing phonetics by computer [Computer program]. Version 6.2.06. Retrieved 14 March 2023 from <https://www.praat.org>
- Bohn O, Flege JE (1992) The production of new and similar vowels by adult German learners of English. *Stud Second Lang Acquis* 14:131–158
- Boudlal A (2001) Constraint interaction in the phonology and morphology of Casablanca Moroccan Arabic. PhD dissertation, Université Mohammed V, Rabat
- Bourhis RY (1982) Language policies and language attitudes: le monde de la francophonie. In: Bouchard R, Gilles H (eds) *Attitudes towards language variation: social and applied contexts*. Edward Arnold, London, 3–4
- Cabrelli Amaro J (2017) The role of prosodic structure in the L2 acquisition of Spanish stop lenition. *Second Lang Res* 33(2):233–269. <https://www.jstor.org/stable/26375880>
- Cabrelli Amaro J, Pichan C (2021) Initial phonological transfer in L3 Brazilian Portuguese and Italian. *Linguist Approaches Biling* 11(2):131–167
- Cardoso W (2007) The development of sC onset clusters in interlanguage: markedness vs. frequency effects. In: Slabakova R, Rothman J, Kempchinsky P, Gavrusseva E (eds) *Proceedings of the 9th generative approaches to second language acquisition conference (GASLA 2007)*, May 18–20, 2007. Cascadia Press, Somerville, pp 15–29
- Coleman J (1999) The nature of vocoids associated with syllabic consonants in Tashlhiyt Berber. In: Ohala JJ, Hasegawa Y, Ohala M, Granville D, Bailey AC (eds) *Proceedings of the 14th international congress of phonetic sciences*. University of California, Berkeley, pp 735–738
- Collins, B., and Mees, I. (Eds.) (2003). *English pronouncing dictionary*. London: Routledge
- Costley T, Reilly C, Gibson H, Kula NC (2023) Teachers' perceptions of teaching in multilingual classrooms: insights from Botswana, Tanzania, and Zambia. In: Karpava S (ed) *Multilingualism, multiculturalism and inclusive education*. Brill, Leiden
- Davidson L (2006) Phonology, phonetics, or frequency: influences on the production of non-native sequences. *J Phon* 34:104–137. <https://doi.org/10.1016/j.wocn.2005.03.004>
- De Angelis G (2005) Interlanguage transfer of function words. *Lang Learn* 55:379–414
- De Angelis G (2007) Third or additional language acquisition. *Multilingual Matters*, Clevedon
- De Angelis G, Selinker L (2001) Interlanguage transfer and competing linguistic systems in the multilingual mind. In: Cenoz J, Hufeisen B, Jessner U (eds) *Cross-linguistic influence in third language acquisition: psycholinguistic perspectives*. Multilingual Matters, Clevedon, pp 42–58
- Dell F, ElMedlaoui M (1988) Syllabic consonants in Berber: some new evidence. *J Afr Lang Linguist* 10:1–17
- Dell F, ElMedlaoui M (2002) *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Kluwer, Dordrecht
- Deterding DH (1997) The formants of monophthong vowels in standard southern British English pronunciation. *J Int Phon Assoc* 27:47–55
- Eckman F (1985) Some theoretical and pedagogical implications of the markedness differential hypothesis. *Stud Second Lang Acquis* 7:289–307. <https://doi.org/10.1017/S0272263100005544>
- Eckman FR, Elreyes A, Iverson GK (2003) Some principles of second language phonology. *Second Lang Res* 19:169–208
- Edmondson WJ (2001) Mehrsprachigkeit: was leistet die schulische Ausbildung? In: Dagmar A-T, Gerhard B (eds) *Mehrsprachiges Europa. Festschrift für Michael Wendt zum 60. Geburtstag*. Narr, Tübingen, pp 131–147
- Ennaji M (2001) De la Diaglossie a la quadriglossie. *Langue et Linguistique* 8:49–64
- Erling EJ, Adinolfi L, Hultgren AK (2017) Multilingual classrooms: opportunities and challenges for English medium instruction in low and middle income contexts. British Council and Education Development Trust, London

- Erling EJ, Clegg J, Rubagumya C, Reilly C (eds) (2021) Multilingual learning and language supportive pedagogies in sub-Saharan Africa. Routledge, London
- Escudero P, Wanrooij K (2010) The effect of L1 orthography on non-native vowel perception. *Lang Speech* 53(3):343–365. <https://doi.org/10.1177/0023830910371447>
- Fairbanks G (1960) Voice and articulation drillbook, 2nd edn. Harper & Row, New York pp 124–139
- Falk Y, Bardel C (2010) The study of the role of the background languages in third language acquisition. The state of the art. *IRAL Int Rev Appl Linguist Lang Teach* 48(2–3):185–219
- Falk Y, Bardel C (2011) Object pronouns in German L3 syntax: evidence for the L2 status factor. *Second Lang Res* 27:59–82
- Flege JE (1995) Second language speech learning: theory, findings and problems. In: Strange W (ed) *Speech perception and linguistic experience: issues in cross-language research*. York Press, Timonium
- Gafos A (2002) A grammar of gestural coordination. *Nat Lang Linguist Theory* 20:269–337
- Gafos A, Hoole PM, Roon K, Zeroual C (2010) Variation in timing and phonological grammar in Moroccan Arabic clusters. In: Fougeron C, Kühnert B, d’Imperio M, Vallé N (eds) *Variation, detail and representation*. Mouton, Berlin/New York, pp 657–698
- Gago Gomez L (2019) Standard Arabic and Moroccan Arabic in the Royal Academy of Morocco Latest Publications (2010–2013). In: Miller C, Barontini A, Germanos M, Guerrero J, Pereira C (eds) *Studies in Arabic dialectology and sociolinguistics: proceedings of 12th Association Internationale de Dialectologie Arabe*. Institut des Recherches et d’Etudes sur les Mondes Arabes et Musulmans (IREMAM), pp 1–10
- Gallardo del Puerto F (2007) Is L3 phonological competence affected by the learner’s level of bilingualism? *Int J Multiling* 4(1):1–16
- Gass S, Selinker L (2008) *Second language acquisition*. Routledge, New York
- Gordon T (2007) *Teaching young children a second language*. Praeger, Westport
- Guion S, Pederson E (2007) Investigating the role of attention in phonetic learning. In: Bohn O-S, Munro M (eds) *Second language speech learning: the role of language experience in speech perception and production*. John Benjamins Publishing, Amsterdam, pp 99–116
- Guion S, Flege JE, Liu S, Yeni-Komshian G (2000) Age of learning effects on the duration of sentences produced in a second language. *Appl Psycholinguist* 21:205–228. <https://doi.org/10.1017/s0142716400002034>
- Gut U (2010) Cross-linguistic influence in L3 phonological acquisition. *Int J Multiling* 7(1):19–38
- Hamann S, Colombo I (2017) A formal account of the interaction of orthography and perception. *Nat Lang Linguist Theory* 35(3):683–714
- Hammarberg B (2001) Roles of L1 and L2 in L3 production and acquisition. In: Cenoz J, Hufeisen B, Jessner U (eds) *Cross-linguistic influence in third language acquisition: psycholinguistic perspectives*. Multilingual Matters, Clevedon, pp 21–41
- Hammarberg B, Hammarberg B (2009) Re-setting the basis of articulation in the acquisition of new languages: a third-language case study. In: Hammarberg B (ed) *Processes in third language acquisition*. Edinburgh University Press, Edinburgh, pp 74–85
- Harris Z (1942) The phonemes of moroccan arabic. *J Am Orient Soc* 62:309–318
- Heath J (1987) *Ablaut and ambiguity: phonology of a Moroccan Arabic dialect*. State University of New York Press, Albany, 336 p
- Heath J (2002) *Jewish and Muslim dialects of Moroccan Arabic*. Curzon, London
- Ingram JCL, Park SG (1997) Inter-language vowel perception and production by Korean and Japanese listeners. *J Phon* 25:343–370
- Jones D (1918) *An outline of english phonetics*. Teubner, Leipzig
- Kamiyama T (2007) Acquisition of French vowels by Japanese-speaking learners: close and close-mid rounded vowels. Paper presented at the L3 phonology satellite workshop of ICPhS XVI, Freiburg, Germany
- Keegan D (1986) Interaction and communication. In: Keegan D (ed) *The foundations of distance education*. Croom Helm, Kent, pp 89–107

- Kenstowicz M, Louriz N (2009) Reverse engineering: emphatic consonants and the adaptation of vowels in French loanwords into Moroccan Arabic. *Brills Ann Afroasiat Lang Linguist* 1(1): 41–74
- Kossmann MG, Stroomer HJ (1997) Berber phonology. In: Kaye AS (ed) *Phonologies of Asia and Africa 2*. Eisenbrauns, Winona Lake, pp 461–475
- Ladefoged P, Johnson K (2011) *A course in phonetics*. Wadsworth, Boston
- Lahrouchi M (2018) Syllable structure and vowel-zero alternations in Moroccan Arabic and Berber. In: Agwuele A, Bodomo A (eds) *The Routledge handbook of African linguistics*. Routledge, London, pp 168–180
- Lindsey G (2012) The British English vowel system. <http://englishspeechservices.com/blog/british-vowels/>. Accessed 10 Mar 2023
- Lipińska D (2015a) Production of L3 vowels: is it possible to separate them from L1 and L2 sounds? *Res Lang* 13(1):77–92
- Lipińska D (2015b) The influence of L2 status on L3 pronunciation. *English Insights* 1:69–85. Retrieved from <http://www.humanitas.edu.pl/resources/upload/dokumenty/Wydawnictwo/English%20Insights/DOROTA%20LIPI%C5%83SKA.pdf>
- Littlewood WT (1994) *Foreign and second language learning: language acquisition research and its implications for the classroom*. Cambridge University Press, Cambridge
- Listeri J, Poch-Olivé D (1987) Phonetic interference in bilingual's learning of a third language. In: *Proceedings of the XIth international congress of phonetic sciences*, vol 5. Academy of Sciences of the Estonian SSR, Austria/Tallinn, pp 134–147
- Llama R, Cardoso W, Collins L (2010) The influence of language distance and language status on the acquisition of L3 phonology. *Int J Multiling* 7:39–57
- Louriz N (2004) *Acquisition of L3 phonology: a case study*. PhD dissertation. Unpublished. University of Essex, UK
- Louriz N (2007) Alignment in L3 phonology. *Langues et Linguistique* 18(19):129–160
- Maas U (2011) *Marokkanisches Arabisch*. LINCOM Sprachlehrbucher, München
- Maas U, Procházka S (2012) Moroccan Arabic in its wider linguistic and social contexts. *STUF Lang Typol Univ* 65(4):329–357
- McKinney C (2020) Decoloniality and language in education: transgressing language boundaries in South Africa. In: Windle JA, de Jesus D, Bartlett L (eds) *The dynamics of language and inequality in education*. Multilingual Matters, Bristol, pp 115–132
- Meisel J (1983) Transfer as a second language strategy. *Lang Commun* 3:11–46
- Ouni S, Laprie Y (2003) A study of the French vowels through the main constriction of the vocal tract using an acoustic-to-articulatory inversion method. In: *International Congress of Phonetic Sciences*, Barcelona
- Parker S (2012) Sonority distance vs. sonority dispersion – a typological survey. In: Parker S, Lahiri A (eds) *The sonority controversy*. Mouton, Berlin, pp 101–166
- Pineda Mora GM, Conklin J, Vera Diettes KJ, Dmitrieva O (2019) Production and perception of English low-mid vowels by speakers of Colombian Spanish in English language immersion environment. In: *Proceedings of the 19th international congress of phonetic sciences*, Melbourne, Australia, pp 3627–3631
- Reilly C, Bagwasi MM, Costley T, Gibson H, Kula NC, Mapunda G, Mwansa J (2022) ‘Languages don’t have bones, so you can just break them’: rethinking multilingualism in education policy and practice in Africa. *J Br Acad* 10(4):1–20
- Reilly C, Gibson H, Kula NC, Costley T, Mapunda G, Mwansa J (In press) Bringing the outside in: attitudes towards multilingual competence in Zambia and Tanzania. In: Mohr S, Ferrara L (eds) *Learning languages, being social*. Anthropological linguistics series. Mouton, London/New York
- Ridouane R (2008) Syllables without vowels: phonetic and phonological evidence from Tashlhiyt Berber. *Phonology* 25(2):321–351
- Ridouane R (2011) Schwa elements in Tashlhiyt word-initial clusters. *Lab Phonol* 2(2):275–300
- Ridouane R (2014) Tashlhiyt Berber. *J Int Phon Assoc* 44(2):207–221. <https://doi.org/10.1017/S0025100313000388>

- Ridouane R (2016) Leading issues in Tashlhiyt phonology. *Lang Linguist Compass* 10(11): 644–606. <https://doi.org/10.1111/lnc3.12211>
- Ringbom H (1987) The role of the first language in foreign language learning. *Multilingual Matters*, Clevedon
- Ringbom H (2002) Levels of transfer from L1 and L2 in L3-acquisition. In: Ytsma J, Hooghiemstra M (eds) *Proceedings of the second international conference on trilingualism*. Fryske Akademie, Leeuwarden
- Rivers WM (1979) Learning a sixth language: an adult learner's daily diary. *Can Mod Lang Rev* 36(1):67–82
- Rothman J (2010) On the typological economy of syntactic transfer: word order and high/low attachment preference in relative clause interpretations in L3 Brazilian Portuguese. *IRAL Int Rev Appl Linguist Lang Teach* 48(2/3):245–273
- Rothman J (2011) L3 syntactic transfer selectivity and typological determinacy: the typological primacy model. *Second Lang Res* 27(1):107–127
- Rothman J, Cabrelli Amaro J (2010) What variables condition syntactic transfer?: a look at the L3 initial state. *Second Lang Res* 26(2):189–218. <https://doi.org/10.1177/0267658309349410>
- Sadiqi F (2001) La dynamique du berbère au Maroc. *Langue et Linguistique* 8:65–79
- Tremblay M-C (2007) L2 influence on L3 pronunciation: native-like VOT in the L3 Japanese of English-French bilinguals. Paper presented at the satellite workshop of ICPHs XVI, Freiburg Germany
- Tremblay MC (2010) Comparing the perceptual abilities of monolinguals, bilinguals and multilinguals: a combined behavioural and event-related potential experiment. PhD dissertation, University of Ottawa
- Trofimovich P, Baker W (2006) Learning second language suprasegmentals: effects of L2 experience on prosody and fluency characteristics on L2 speech. *Stud Second Lang Acquis* 28:1–30
- Wang Y, Jongman A, Sereno J (2003) Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *J Acoust Soc Am* 113:1033. <https://doi.org/10.1121/1.1531176>
- Weckwerth J (2021) Vowels. In: Knight R, Setter J (eds) *The Cambridge handbook of phonetics*. Cambridge University Press, Cambridge, pp 40–64
- Werker JF (1986) The effect of multilingualism on phonetic perceptual flexibility. *Appl Psycholinguist* 7:141–155. <https://doi.org/10.1017/S0142716400007360>
- Williams S, Hammarberg B (1998) Language switches in L3 production: implications for a polyglot speaking model. *Appl Linguis* 19:295–333
- Wilson C, Davidson L (2009) Bayesian analysis of non-native cluster production. In: *Proceedings of the NELS 40*. MIT, Cambridge, MA
- Wrembel M (2010) L2-accented speech in L3 production. *Int J Multiling* 7(1):75–90
- Wrembel M (2011) Cross-linguistic influence in third language acquisition of voice onset time. In: Lee WS, Zee E (eds) *Proceedings of the 17th international congress of phonetic sciences*, pp 17–21
- Wrembel M (2013) Foreign accent ratings in third language acquisition: the case of L3 French. In: Waniek-Klimczak E, Shockey L (eds) *Teaching and researching English accents in native and non-native speakers*. Springer, Berlin, pp 31–47
- Wulff S, Ellis CN (2018) Usage-based approaches to second language acquisition. In: Miller DT, Bayram F, Rothman J, Serratrice L (eds) *Bilingual cognition and language: the state of the science across its subfields*. John Benjamins Publishing, Amsterdam, pp 37–56

Glide: High Vowel Alternation at the Syntax–Phonology Interface

4

Mohamed Lahrouchi and Rachid Ridouane

Contents

Introduction	87
Glides Versus High Vowels: Features, Representations and Alternations	89
Amazigh Phonemic System: An Outline	89
Glides Versus High Vowels: Features and Representations	90
Glide – High Vowel Alternations in Amazigh	90
Boundaries, Phases, and Spell-Out: The Basics	94
Grammatical Carriers in Phonology	94
Derivation by Phase	94
Phasal Spell-Out of the High Vocoids I, U	95
Emphasis Spread in Verbs and Nouns: Acoustic Data	97
Conclusion	99
References	100

Introduction

Anyone who has once thumbed through a textbook on Amazigh (Berber) phonology has probably come across one of the most productive processes in the language, namely, the alternation of glides and high vowels, some aspects of which still challenge standard phonological theories. In Tashlhiyt, like in many other Amazigh varieties, glides typically appear in the immediate vicinity of a vowel, in complementary distribution with the corresponding high vowels, as the following examples illustrate (the alternating segments are in bold):

M. Lahrouchi (✉)

CNRS & University Paris 8, Paris, France

e-mail: mlahrouchi@univ-paris8.fr

R. Ridouane

Laboratoire de Phonétique et Phonologie, CNRS & Sorbonne Nouvelle, Paris, France

e-mail: rachid.ridouane@sorbonne-nouvelle.fr

(1)

gru	“pick up”	agraw	“assembly”
bri	“crush”	abraj	“crushed seeds”
nu	“be cooked”	tinwi	“cooking”

Standard analyses generally address this issue from a syllabic point of view, claiming that glides and high vowels are phonetic reflexes of two underlying high vocoids: I and U. These vocoids surface as high vowels when attached to a nucleus position, and as glides when associated to margin positions (cf. Bensoukas 2001: 35; Boukhris et al. 2008: 23; Dell and Elmedlaoui 2002: 190–198; Iazzi 2018: 429; Lahrouchi 2001, 2010, 2013; Lahrouchi and Ségéral 2010; Lahrouchi and Kern 2018: 504; Ridouane 2014: 214, among others).

The problem with the above generalization arises when considering contexts where the underlying vocoids surface as high vowels when they should normally be realized as glides. Let us consider the dative forms of the verbs *gru* and *bri*, compared to those of *skr* “do” and *fk* “give”:

(2)

	<i>Verb</i>	<i>Dative</i>
a.	<i>gru</i>	<i>gru-j-as</i>
	<i>bri</i>	<i>bri-j-as</i>
b.	<i>skr</i>	<i>skr-as</i>
	<i>fk</i>	<i>fk-as</i>

From the examples in (2b) one concludes that the phonological form of the dative morpheme is /as/, immediately following the final consonant of the verb base. When attached to verbs like in (2a), the initial vowel of the dative marker should normally allow the verbs final vocoid to surface as a glide, leading to the forms **grwas* and **brjas*. Instead, the high vocoids are realized as vowels, resulting in a hiatus context, which is then resolved by means of *j*-epenthesis.

The question then arises as to whether this type of phonological opacity (Kiparsky 1976) can be addressed without having recourse to any specific boundaries, junctures (Trager 1962; Lehiste 1965), phonological cycles (Kiparsky 1982), or any other morpheme-specific phonology (Pater 2010).

The key to understanding this phenomenon lies, we argue, in the morphosyntactic structure of the forms at stake. In line with recent works on the syntax–phonology interface (cf. Marvin 2002; Marantz 2007; Newell and Piggott 2014; Samuels 2010; among others), we claim that the behavior of the high vocoids in the forms of the type in (1) and (2) is the result of the application of spell-out and phase impenetrability condition (Chomsky 2001) at different levels in the syntactic structure. In particular, we show that *vP* corresponds to a phase where I and U are spelled out as high vowels before the dative enclitic /-as/ is added. The hiatus is then resolved by means of *j*-epenthesis, leading to *grujas* and *brijas*. In *agraw* and *abraj*, U and I surface as glides since they belong to the same phase (*nP*) as the neighboring vowel /a/.

External evidence in favor of this analysis will be drawn from the phenomenon of emphasis spread, also called “dorsopharyngealization” and widely attested in many Afroasiatic languages, including Amazigh (cf. Elmedlaoui 1985, 1995; Boukous 1990; Lasri 1991; Ridouane 2003, among others). It will be shown that in Amazigh the emphatic coronals spread their feature to the neighboring segments within the vP and nP domains. That is, verbal and nominal bases containing an emphatic consonant will be entirely emphaticized to the exclusion of their suffixes, which will remain unaffected.

The chapter is structured as follows: Section “[Glides Versus High Vowels: Features, Representations and Alternations](#)” presents the data more fully, and illustrates the set of alternations just outlined. It also explains how the standard syllabic approach to these alternations remains valid, provided that extra-phonological information is taken into account. The examples in this chapter are all from the Tashlhiyt variety, but the facts are similar in other Amazigh varieties (cf. Destaing 1920; Basset 1952; Appelgate 1970; Guerssel 1986; Kossmann and Stroomer 1997: 467; Iazzi 2018: 429; among others). Section “[Boundaries, Phases, and Spell-Out: The Basics](#)” provides the theoretical background necessary for understanding the proposed analysis. Our phase-based analysis is presented in section “[Phasal spell-Out of the High Vocoids I, U](#)”. Section “[Emphasis Spread in Verbs and Nouns: Acoustic Data](#)” provides another piece of evidence in favor of our analysis: It will be shown that emphasis spreads within the domains of nP and vP, exactly where glides alternate with high vowels. Section “[Conclusion](#)” concludes the chapter.

Glides Versus High Vowels: Features, Representations and Alternations

Amazigh Phonemic System: An Outline

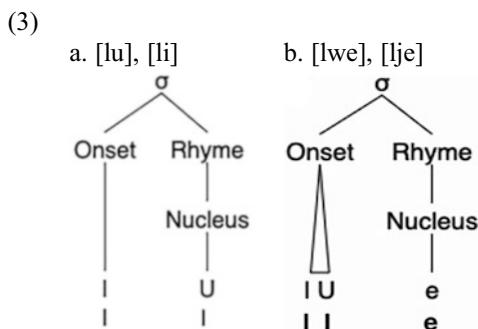
Tashlhiyt, like almost all other Amazigh varieties, has a simple vocalic system with three phonemic vowels /i, a, u/, and a central short vowel, generally referred to as schwa [ə], whose phonological status is highly controversial. According to many scholars (Dell and Elmedlaoui 2002; Ridouane 2008; among others), this short vowel, found in certain consonant clusters, is a mere transitional vocoid that has no syllabic status (*contra*. Coleman 1996, 2001; Puech and Louali 1999).

As for consonants, Tashlhiyt has 35 phonemic units (including the glides /j, w/), each of which has a geminate counterpart. Although the phonemic status of the two glides is clearly established, as they contrast underlyingly with high vowels (e.g., *rwl* “escape” / *ruħ* “go away,” *r’mi* “be tired” / *rwi* “soil”), and like any other consonant they undergo imperfective gemination (e.g., *lkkm* “arrive.imperf” / *rwwi* “soil.imperf”), they very often appear in complementary distribution with the corresponding high vowels (further discussion and data in section “[Glide – High Vowel Alternations in Amazigh](#)”). This fact has led many linguists to claim that glides and high vowels are merely phonetic reflexes of the same underlying segments, whose syllabic representations can easily account for their surface realization.

Glides Versus High Vowels: Features and Representations

Glides, also termed semi-consonants or semi-vowels, have been largely documented in the literature, from both empirical and theoretical points of view. Their behavior has constantly posed a problem to phonological theory. On the one hand, they closely interact with vowels, alternating with them in various well-defined contexts – not only in the language under scrutiny, but also in many other typologically unrelated languages (see Rosenthal 1994; Padgett 2008; Levi 2011). On the other hand, they behave as consonants, underlyingly contrasting with vowels.

With the advent of syllable theory, many phonologists have proposed that the difference between vowels and glides can be derived directly from syllable structure (cf. Clements and Keyser 1983; Kaye and Lowenstamm 1984; Levin 1985; Selkirk 1982, 1984), thus making redundant the standard SPE featural distinctions: [\pm vocalic], [\pm consonantal], [\pm syllabic] (Chomsky and Halle 1968: 303, 354), and [\pm approximant] (Clements 1990: 293). That is, vowels appear in the syllable nucleus, whereas glides attach to coda and onset positions. In this respect, Kaye and Lowenstamm (1984: 130–138) claimed that in French the high vocoids /U/ and /I/ are realized as vowels in verbs like *loue* “rent.IMPERATIVE.SG” and *lie* “tie.IMPERATIVE.SG,” and as glides in the corresponding infinitive forms *louer* and *lier*. This is illustrated in the representations in (3).



The underlying segments /U, I/ behave as syllabic in (3a), while they form part of a complex onset in (3b). Their position in the syllable structure allows assigning to them a phonetic interpretation without the mediation of any specific feature. The same reasoning underlies the representation of alternating glides and high vowels in Amazigh, as it will be shown in the next section.

Glide – High Vowel Alternations in Amazigh

This section provides the empirical basis for the generalizations briefly introduced in the previous sections, through a closer, fuller examination of the glide-high vowel alternations in Tashlhiyt.

Contexts of Alternation

The paradigm in (4) elaborates on (1) above. It shows the structural environments where high vowels alternate with glides.

(4)

	<i>Verb</i>		<i>Noun</i>	
a.	gru	“pick up”	agraw	“assembly”
	χlu	“destroy”	amχlaw	“madman”
	aru	“give birth to”	arraw	“child”
	nu	“be cooked”	tinwi	“cooking”
b.	bri	“crush”	abraj	“crushed seeds”
	sti	“sort”	astaj	“sorting”
	fsi	“untie”	afssaj	“untying”
	ngi	“overflow”	angaj	“overflowing”
	mlilli	“feel dizzy”	timlillaj	“dizziness”

Each pair in these examples clearly shows how the verb’s final vowel turns into a glide when, for morphological reasons, a vowel appears in its immediate vicinity: this is especially the case for the infix -a-, which is inserted after the stem’s second consonant (and which is similarly found in derivations like *skr* “to dry” > *askar* “wood, tree,” *frs* “be sharp” > *afras* “sharpening”). This infix allows [u] to change into [w] (4a), and [i] into [j] (4b). The affixal vowel can also appear after the target segment like in the nominal form *tinwi* where [i] turns the stem’s final [u] into [w].

Similar alternations are found in plural nouns, as shown in the examples in (5).

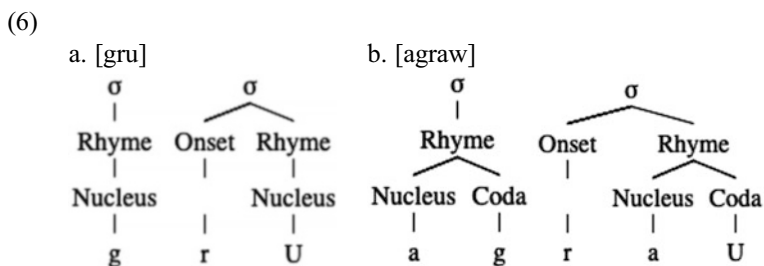
(5)

<i>SG</i>	<i>PL</i>	
iflu	iflwan	“plank”
ikru	ikrwan	“kid”
amdlu	imdlawn	“cloud”

The plural suffix -an (also found in forms like *ifrg-an* “enclosure” (< *sg ifrg*) and *ifr-an* “caves” (< *sg ifri*)) is responsible for the semi-vocalization of final [u] of the singular form. Many other formations exhibit similar patterns in Amazigh, including the construct state forms whose vowel marker [u-] alternates with [w-] (e.g., u-frux “boy.cs” versus w-adgal “widower.cs”; cf. Dell and Elmedlaoui 2002; Lahrouchi 2010, 2013; among others) and the perfective forms where the 3MS marker [i-] alternates with [j-] (e.g., i-skr “he did” vs. j-ukr “he stole”). The focus on the forms of the type in (4) and (5) will become more evident when analyzing their datives and demonstratives, whose final high vowels resist alternation. The basic rule underlying all of the examples examined so far is the avoidance of any hiatus context: semi-vocalization is one of these strategies that Amazigh utilizes in this respect (In Amazigh, hiatus may also be resolved by means of j-epenthesis, a fact we will be highlighting in subsequent sections (e.g., *fk-as* “give him” vs. ini-j-as “tell him”), or

by deletion of one of the vowels in contact (e.g., afus-**inu** “my hand” vs. tasa-**nu** “my liver, my dear,” azɛɾɛu (SG) vs. izɛɾɛ-an (PL) “stone,” itri (SG) vs. itr-an (PL) “star”). The reader is referred to Boukous (1979, 2009) and Iazzi (2018: 459), for further details and analysis.).

Many Amazigh phonologists have analyzed these alternating segments as variants of the same underlying segments, whose realizations are fully derivable from their position in the syllable structure (cf. Dell and Elmedlaoui 1985, 2002; Bensoukas 1994, 2001; Lahrouchi 2001, 2013; Lahrouchi and Ségéral 2010; Soutsane 2008; Ridouane 2014; among others). In line with standard approaches to syllabic structure (cf. Pike and Pike 1947; Levin 1985; Blevins 1995; see also Bosch 2011; Goldsmith 2011; Scheer 2015, for detailed reviews of the relevant literature), the representation of the verb *gru* “pick up” and its nominal form *agraw* “assembly” is shown in (6).



A thorough discussion of the above syllabic representations, and the ability for any segment to be syllabic in Tashlhiyt, would take us beyond the scope of this chapter and rates a further discussion elsewhere (cf. Dell and Elmedlaoui 1985, 2002; Boukous 1987; Lahrouchi 2018). We simply draw attention to the syllabic positions where the vocoid /U/ appears: when attached to a nucleus position, the vocoid is realized as a high vowel (6a), while it surfaces as a glide when associated with a margin position (6b). Similar representations hold for the vocoid /I/ in words like *bri* “crush” and *abraj* “crushed seeds.” We refer to Dell and Elmedlaoui (2002: 198) for similar syllabic representations of the high vocoids /I/ and /U/ in words like *t-suj* “she let pass,” and *t-zwi* “she beat down.”

The syllabic approach entirely accounts for the high vowel-glide alternations that occur within a specific domain, which the present study aims to define in a principled way. The following section turns to cases where the so far called “alternating high vowels” cease to undergo alternation.

Lack of Alternation

One striking context in which the high vowels of the forms in (4) resist alternation arises with the dative and 3rd person object suffix /-as/. This is shown in the following examples.

(7)

	<i>Verb</i>		<i>Dative</i>	
a.	gru	“pick up”	gru-j-as	“pick to him/her”
	χlu	“destroy”	χlu-j-as	“destroy for him/her”
	kru	“rent”	kru-j-as	“rent him/her”
	zru	“delouse”	zru-j-as	“delouse him/her”
b.	bri	“crush”	bri-j-as	“crush for him/her”
	sti	“sort”	sti-j-as	“chose for him/her”
	fsi	“untie”	fsi-j-as	“untie him/her”
	zʕli	“isolate”	zʕli-j-as	“put aside for him/her”

Unlike in (4), the verbs’ final vowels resist gliding even though they are immediately followed by a vowel-initial suffix. In such a context, the high vowels should normally surface as glides, leading to the dative forms **grwas* and **brjas*, to take but two examples. Instead, a glide [j] is inserted in order to avoid hiatus.

The same phenomenon takes place in the nominal forms of the type in (5). The high vowels of their singular forms resist alternation when immediately preceding the demonstrative enclitic /-ad/, as shown in (8).

(8)

<i>Noun</i>		<i>Noun + Demonstrative</i>	
iflu	“plank”	iflu-j-ad	“this plank”
ikru	“kid”	ikru-j-ad	“this kid”
amdlu	“cloud”	amdlu-j-ad	“this cloud”

Based on similar facts in Tamazight, Guerssel (1986: 103) claimed that the differences observed between non-alternating high vowels as in *t-uri* “she wrote” versus *t-uri-j-ax* “she wrote us” (not **turjax*), and *t-ssu* “she made a bed” versus *t-ssu-j-ax* “she made us a bed” (not **tsswax*), and alternating ones like in *i-ru* “he cried” versus *aha j-ru* “then he cried” receive “a natural explanation if a phonemic distinction between glides and high vowels is established.”

This has been proposed in the case of Tamazight. The data under scrutiny hardly fit into such an analysis for the very simple reason that the final high vowels in (7) and (8) actually undergo alternation within the stem domain, as shown previously in examples in (4) and (5). They resist alternation only when followed by certain vowel initial suffixes, such as the dative *-as* and the demonstrative *-ad*. While keeping a syllabic approach to these alternations, we will call on those formal mechanisms, which integrate extra-phonological information capable of explaining why alternating segments cease to do so beyond a certain domain. Cyclic derivation, at the interface between syntax and phonology, nowadays referred as to “phases,” is one such device that better accounts for these paradoxical cases where the same underlying segments may or may not undergo alternation. Before expanding on this issue, we provide some theoretical background necessary to understanding our analysis.

Boundaries, Phases, and Spell-Out: The Basics

Grammatical Carriers in Phonology

One undeniable property that characterizes the contexts in which high vowels do not alternate with glides is that they contain a boundary between the stem and the enclitic (dative *-as* and demonstrative *-ad*). The question is what the exact nature of these boundaries is and how they affect the phonological interpretation of sound strings: do they carry any morphosyntactic information which impacts phonology?

Many studies have long been investigating various phenomena whereby automatic phonological processes fail to occur when they come up against certain type of boundaries. In the 1960s, structuralist phonologists (cf. Trager 1962; Lehiste 1965) drew a distinction between phonemes as basic units of phonological analysis, and juncture phonemes as carriers of grammatical information in phonology. Generative phonologists resorted to similar objects, including classical SPE types of diacritics (*#*, and the like) and prosodic constituents (*ω*, *φ*, and the like) as means by which phonologically relevant domains are delimited. These so-called representational ways of implementing morphosyntactic information in phonology are distinguished from derivational solutions, successively known as transformational cycles, phonological cycles, levels, or strata (Mascaró 1976; Kiparsky 1982; Mohanan 1982; Bermúdez-Otero 2011), and phases (Chomsky 2001, 2008). The reader is referred to Scheer (2011) for a thorough review of the relevant literature.

Derivation by Phase

Derivation by phase allows for a model of grammar in which pre-specified chunks of syntactic structure are sent to the phonological and semantic components. According to Chomsky (2001, 2008), *vP* and *CP* correspond to phases within which lexical material is inserted and constituents may move up to higher syntactic positions. Generalizing from *vP*, Marantz (2001, 2007) claims that any category-forming projection uniformly defines a phase within which the phonological and semantic properties of words are interpreted.

Relying on these proposals, many studies have addressed various non-automatic phonological processes that standard analyses fail to capture in a satisfactory manner. In her fundamental work, Marvin (2002) addressed among other phenomena, the opacity characterizing schwa-insertion in words like *meter* [mi:tər], *metering* [mi:təriŋ], and *metric* [metrik]. She argued that the adjectival suffix *-ic*, which prevents schwa from appearing in *metric*, is spelled out in the same phase as /metr/, as opposed to the gerund suffix *-ing*, which is added later in the syntactic structure, once /metr/ is spelled out along with its schwa (see also

Marvin 2013, for a phase-based account for stress assignment in English). The Phase Impenetrability Condition (Chomsky 2001) facilitates an explanation as to why a previously spelled out schwa resists deletion. Various phenomena in languages like German (Kratzer and Selkirk 2007), Basque (Samuels 2010), Ojibwe (Newell and Piggott 2014), Abruzzese (D’Alessandro and Scheer 2015), and Kiowa (Miller 2018) have since been analyzed in a comparable manner despite differences in the versions of phase theory and a direct versus indirect reference to morphosyntactic categories (cf. Scheer 2008, 2011). What these studies share as common is that morphosyntactic chunks that correspond to phases locally determine the domain of application of certain phonological processes.

Unlike phonological cycles that may appear as ad hoc stipulations, and that lack any external evidence, phases have independent morphosyntactic motivation outside the realm of phonology. Moreover, derivation by phase allows specific chunks to be spelled out, thus leading to a grammar whose computation is simpler in terms of memory load and processing.

In the following section, we return to glide-high vowel alternations. We argue that phasal spell out better explains why high vowels fail to alternate with glides when followed by certain vowel-initial suffixes.

Phasal Spell-Out of the High Vocoids I, U

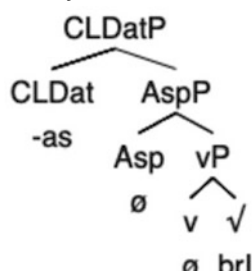
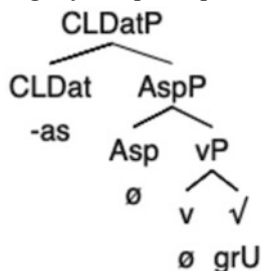
As we have seen in section [Glide – High Vowel Alternations in Amazigh](#), final high vowels resist alternation with glides in verbs like *bri* “crush” and *gru* “pick up.” When adjacent to the vowel-initial dative enclitic /-as/, [i] and [u] should normally surface as glides; especially since they do so in the corresponding nominal forms *abraj* “crushed seeds” and *agraw* “assembly.” Instead, a glide [j] is inserted between the verb’s final vowel and the enclitic in order to resolve hiatus, leading to *brijas* “crush him/her” and *grujas* “pick for him/her.” Similarly, the final high vowels of singular nouns like *iflu* “plank” and *ikru* “kid,” which surface as glides in the corresponding plural forms *iflw-an* and *ikrw-an*, paradoxically remain unchanged when followed by the demonstrative suffix /-ad/.

The key to understanding these opaque cases lies in the syntactic structure of the dative and demonstrative forms. The failure of the high vowels to change into glides in the dative forms receives a natural explanation once we assume that the verbal projection (vP) constitutes a phase within which the phonological and semantic features of the verb base are fully interpreted. More specifically, the high vocoids of verbs like *gru* and *bri* are spelled out inside the vP phase as vowels before the dative enclitic /-as/ is added. The phase domain within which these verbs are interpreted is rendered opaque to further operations, including the suffixation of the dative enclitic whose initial vowel cannot recover the phonological identity of the verb’s final segment: hence, j-epenthesis. This is illustrated in the structures in (9a) and (9b),

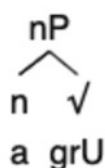
as opposed to the ones in (9c) and (9d) where the same vocoids undergo semi-vocalization in the corresponding nominal forms.

(9)

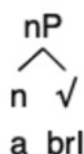
a. *gru-j-as* “pick up for him!” b. *bri-j-as* “crush for him!”



c. *agraw* “assembly”



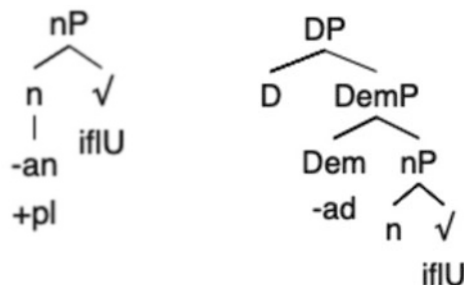
d. *abraj* “crushed seeds”



These structures are simplified for expository reasons. The reader is referred to Ouhalla (1988), Ouali (2011), Achab (2012), among others, for a thorough treatment of verb and clause structure in Amazigh. Of immediate concern is that the dative clitic is base generated outside the vP domain, as the structures in (9a, b) illustrate (cf. Ouali 2005, 2011; Omari 2012, on this issue). Under the assumption that Phase Impenetrability Condition holds in phonology, the phonological string spelled out within the vP domain becomes inaccessible to subsequent operations, leading to a situation where the high vocoids I and U are realized as vowels before the dative clitic is added. The resulting hiatus is then resolved by means of j-epenthesis. In other words, the rule regulating glide-high vowel alternations holds only within the same domain/phase: The high vocoids I and U surface as glides when adjacent to a vowel, as in the forms represented in (9c) and (9d), and as vowels elsewhere.

The same reasoning holds for nouns like *iflu* “plank” and *ikru* “kid,” previously introduced in (5) and (8): Their high vowel turns into a glide when followed by the plural suffix *-an*, while it remains unchanged when followed by the demonstrative clitic *-ad*. The reason, we argue, is that the high vocoid U is spelled out along with the plural suffix within the same phase (nP). Conversely, the demonstrative marker is generated higher in the structure once the high vocoid is spelled out as a vowel. The representations in (10) illustrate both situations.

(10)

a. *iflw-an* “plank.pl” b. *iflu-j-ad* “this plank”

Two observations are in order. First, number and plurality are assumed to be realized as a feature under the head of noun phrase, as shown in (10a) (see Lowenstamm 2008; Acquaviva 2008; and Kramer 2012). Second, the demonstrative marker *-ad* heads its own projection below DP, as shown in (10b). The linear order (postposed to the noun) is derived through movement of N to the D position. The reader is referred to Belkadi (2017: 117) for further details and references, and to El Moujahid (1997: 223) for an alternative analysis according to which the demonstrative acts as a complement of N.

The crucial point about the above structures is that the high vocoid U surfaces as a glide in (10a) since it is spelled out along with the plural marker within the nP phase. Conversely, it surfaces as a vowel in (10b) because it is spelled out before the vowel-initial suffix *-ad/* is added. The resulting hiatus is then resolved by means of j-epenthesis, exactly like in (9a) and (9b).

External evidence in favor of this analysis is drawn from emphasis spread. In the next section, we show that phasal spell-out better explains how the verbal and nominal bases, containing an emphatic consonant, are entirely pharyngealized to the exclusion of any segmental material generated outside the domains of vP and nP.

Emphasis Spread in Verbs and Nouns: Acoustic Data

Emphasis, also called dorsopharyngealization, refers to a secondary articulation whereby a specific set of Tashlhiyt consonants is produced with a backward movement of the tongue toward the posterior pharyngeal wall, while the anterior part of the tongue is lowered (Ridouane 2009). At the phonemic level, emphasis is a property of individual segments. Tashlhiyt includes the following set of emphatic consonants, all of which are coronal: /tʰ, tːʰ, dʰ, dːʰ, sʰ, sːʰ, zʰ, zːʰ, ʒʰ, ʒːʰ, rʰ, rːʰ, lʰ, lːʰ/ (El Moujahid 1979; Elmedlaoui 1985, 1995; Boukous 1987, 1990; Lasri 1991; Dell and Elmedlaoui 2002).

At the surface level, dorsopharyngealization is a property that can be displayed by any segment; an instance of emphasis spread by which underlying pharyngealized coronals spread their secondary articulation onto neighboring sounds. For instance, the form /matʕiʃa/ “tomato” contains one underlying emphatic coronal /tʕ/, but all the segments contained in the word can be phonetically dorsopharyngealized: [mʕaʕtʕiʕʃaʕ]. Pharyngealized words may be composed of emphatic coronals only (e.g., [sʕrʕdʕ] “to file a complaint”). In such items there is no principled way of determining whether only one or more coronals are underlyingly dorsopharyngealized.

The exact delimitation of the propagation of dorsopharyngealization is not clearly established. The segments in a CV sequence must be both plain or both emphatic. This requirement must be satisfied regardless of the morphosyntactic relationship between C and V (Dell and Elmedlaoui 2002). But it is generally considered that the domain of this propagation may be larger than the word (Elmedlaoui 1985, 1995; Boukous 1990; Lasri 1991). There are important acoustic differences between pharyngealized coronals and their plain counterparts, which induce highly salient auditory differences between items containing emphatic consonants, and items containing plain ones. The most important and consistent difference is manifested by a lowering of F2 of the vowels contained in the emphatic items. This pattern has been extensively shown in different Arabic varieties, including Moroccan Arabic (Yeou 1997; Zeroual 2000; Lahrouchi and Ridouane 2016). The same pattern has been observed in Tashlhiyt. Ridouane (2003), based on vowel productions in the context of /tʕ/, and its plain counterpart /t/ (e.g., [atʕa] vs. [ata]) by five Tashlhiyt speakers, showed that emphasis lowers F2 of the adjacent vowels in a systematic way. In this context the realization of the vowel /a/ is close to back [ɑ].

We conducted a pilot acoustic study in order to determine whether the vowel /a/ in the dative enclitic /-as/, and the demonstrative /-ad/, is affected by the presence of emphatic consonants within the same words. The acoustic data were recorded by two subjects, producing six pairs of words contrasting emphatic to plain consonants in verbal and nominal forms. Each pair was repeated three times by each speaker. The data set is presented in (11) below. We measured F2 values at the midpoint of the vowel /a/ (bolded in the data below).

(11)

List of items used in the acoustic experiment

Verbs

bdu-j-as	“begin”	bdʕu-j-as	“divide”
ndu-j-as	“demolish”	ndʕu-j-as	“jump”
mdi-j-as	“trap, intercept”	mdʕi-j-as	“taste”

Noun

ssif-ad	“sword”	ssʕif-ad	“summer”
adan-ad	“bowls”	adʕan-ad	“nights”
izri-j-ad	“mugwort”	izʕri-j-ad	“sight”

From our perspective, segmental material generated outside the domains of vP and nP will not be affected by emphasis spread. According to this view, the vowel /a/ in /-as/ and /-ad/ will have similar F2 values, whether the base contains an emphatic

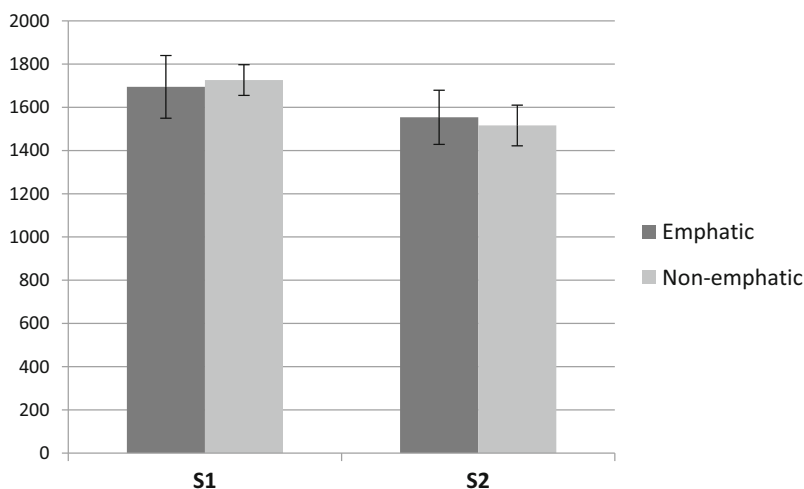


Fig. 4.1 F2 values (Hz) of the vowel /a/ in emphatic and nonemphatic contexts: for the two subjects S1 and S2

consonant, or not. That is, the vowel /a/ in the words /bdu-j-as/ and /bd^hu-j-as/, for example, will display similar F2 values, although the latter word has an emphatic consonant. This is exactly what the results of our study show. The vowel /a/ displays virtually identical F2 values in both the emphatic and nonemphatic contexts. This pattern is illustrated in Fig. 4.1.

The reason why the vowel of the dative and demonstrative suffixes resist emphasis is the same which, from our point of view, prevents the stem-final vowel from surfacing as a glide as previously discussed in section “[Phasal Spell-Out of the High Vowoids I, U](#)”. In other words, vP and nP demarcate a phase, wherein emphasis spreads and glides alternate with the corresponding high vowels.

Conclusion

In this study, we have accounted for glide-high vowel alternations in Amazigh at the interface between syntax and phonology. This phenomenon, which has long resisted exclusively phonological analyses, has been argued to occur within specific syntactic domains, corresponding to phases (vP and nP) wherein I and U in words like *bri* “scratch” and *iflu* “plank” are spelled out as high vowels. Interestingly, these vowels remain unchanged when followed by vowel-initial suffixes generated outside the phasal domain, including the dative /-as/ and the demonstrative /-ad/. Emphasis spread was used as evidence in support of our analysis of glide-high vowel alternation. Based on acoustic data from two subjects, we showed that the dative and demonstrative suffixes, which are generated outside vP and nP domains, remain unaffected by emphasis spread, whether their base contains an emphatic consonant or not.

References

- Achab K (2012) Internal structure of verb meaning: a study of verbs in Tamazight (Berber). Cambridge Scholars Publishing, Newcastle upon Tyne
- Acquaviva P (2008) Lexical plurals: a morphosemantic approach. Oxford University Press, Oxford
- Appelgate J (1970) The Berber Languages. In: Sebeok TA, Ferguson CA, Hodge CT, Paper H, Krueger JR, Schram GM (eds) Current trends in linguistics 6: linguistics in south West Asia and North Africa. Mouton, Berlin, pp 586–661
- Basset A (1952) La langue berbère. In: Handbook of African language 1. Oxford University Press, Oxford
- Belkadi A (2017) Pronominal and clitics systems in Taqbaylit Berber. Pro Quest LLC, Ann Arbor
- Bensoukas K (1994) Tashlhit agentive nouns: an optimality-theoretic approach. D.E.S thesis, Mohammed V University, Faculty of Letters, Rabat
- Bensoukas K (2001) Stem forms in non-templatic morphology of Berber. Doctorat d'Etat thesis, Mohammed V University, Rabat
- Bermúdez-Otero R (2011) Cyclicity. In: van Oostendorp M, Ewen C, Hume E, Rice K (eds) The Blackwell companion to phonology. Wiley-Blackwell, Hoboken, pp 2019–2048
- Blevins J (1995) The syllable in phonological theory. In: Goldsmith J (ed) The handbook of phonological theory. Blackwell, Hoboken, pp 206–244
- Bosch A (2011) Syllable-internal structure. In: van Oostendorp M, Ewen C, Hume E, Rice K (eds) The Blackwell companion to phonology. Wiley-Blackwell, Hoboken, pp 781–798
- Boukhris F, Boumalk A, El Moujahid E, Souifi H (2008) La nouvelle grammaire de l'amazighe. Publications de l'Institut Royal de la Culture Amazighe, Rabat
- Boukous A (1979) La formation du glide en Tamazight (dialecte tachelhit). Traces 2:3–51
- Boukous A (1987) Phonotactique et domaines prosodiques en berbère. PhD dissertation, University Paris 8
- Boukous A (1990) Pharyngalisation et domaines prosodiques en berbère. In: Langue et Littérature. Publications de la Faculté des Lettres et des Sciences Humaines, Sorbonne, pp 153–176
- Boukous A (2009) Phonologie de l'amazighe. Publications de l'IRCAM, Paris
- Chomsky N (2001) Derivation by phase. In: Kenstowicz M (ed) Ken Hale: a life in language. MIT Press, Cambridge, MA, pp 1–52
- Chomsky N (2008) On phases. *Curr Stud Linguist Ser* 45:133
- Chomsky N, Halle M (1968) The sound pattern of English. Harper & Row Publishers, New York
- Clements GN (1990) The role of the sonority cycle in core syllabification. In: Kingston J, Beckman ME (eds) Papers in laboratory phonology I: between the grammar and physics of speech. Cambridge University Press, Cambridge, pp 283–333
- Clements G, Keyser J (1983) CV phonology: a generative theory of the syllable. MIT Press, Cambridge
- Coleman J (1996) Declarative syllabification in Tashlhiyt Berber. In: Durand J, Laks B (eds) Current trends in phonology: models and methods. European Studies Research Institute, University of Salford, Salford, pp 177–218
- Coleman J (2001) The phonetics and phonology of Tashlhiyt Berber syllabic consonants. *Trans Philol Soc* 99:29–64
- D'Alessandro R, Scheer T (2015) Modular PIC. *Linguist Inq* 46(4):593–624
- Dell F, Elmedlaoui M (1985) Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *J Afr Lang Linguist* 7:105–130
- Dell F, Elmedlaoui M (2002) Syllables in Tashlhiyt Berber and in Moroccan Arabic. Kluwer, Amsterdam
- Destaing E (1920) Etude sur le dialecte berbère des Ait Seghrouchen. Editions Ernest Leroux, Paris
- El Moujahid E (1979) Présentation des phonèmes de la langue tamazight, la tachelhiyt d'Igherm Souss. Traces 2:52–78
- El Moujahid E (1997) Grammaire générative du berbère, morphologie et syntaxe du nom en tachelhit. Publications de la Faculté des Lettres et des Sciences Humaines, Sorbonne

- Elmedlaoui M (1985) *Le parler berbère chleuh d’Imdlawn Maroc: Segments et syllabation*. Thèse de Doctorat, Université Paris 8
- Elmedlaoui M (1995) *Aspects des représentations phonologiques dans certaines langues chamito-sémitiques*. Université Mohammed V, Publications de la Faculté des Lettres et des Sciences Humaines, Rabat
- Goldsmith J (2011) Syllables. In: Goldsmith J, Riggle J, Yu A (eds) *The handbook of phonological theory*, 2nd edn. Wiley-Blackwell, Hoboken, pp 162–196
- Guerssel M (1986) Glides in Berber and syllabicity. *Linguist Inq* 17(1):1–12
- Iazzi E (2018) *Norme et variations en amazighe marocain (aspects morphophonologiques)*. Doctorat d’Etat thesis, Ibn Zohr University, Agadir
- Kaye J, Lowenstamm J (1984) De la syllabacité. In: Dell F, Hirst D, Vergnaud JR (eds) *Forme Sonore du Langage: Structure des Représentations en Phonologie*. Hermann, Paris, pp 123–159
- Kiparsky P (1976) Abstractness, opacity, and global rules. In: Koutsoudas A (ed) *The application and ordering of grammatical rules*. Mouton, Berlin
- Kiparsky P (1982) From cyclic to lexical phonology. In: van der Hulst H, Smith N (eds) *The structure of phonological representations*, vol I. Foris, Berlin, pp 131–175
- Kossmann M, Stroomer H (1997) Berber phonology. In: Kaye AS (ed) *Phonologies of Asia and Africa 2*. Eisenbrauns, Ann Arbor, pp 461–475
- Kramer R (2012) A split analysis of plurality: evidence from Amharic. In: Arnett N, Bennet R (eds) *The proceedings of WCCFL*, vol 30. Cascadilla, Somerville, pp 226–236
- Kratzer A, Selkirk E (2007) Phase theory and prosodic spellout: the case of verbs. *Linguist Rev* 24(2–3):93–135
- Lahrouchi M (2001) *Aspects morpho-phonologiques de la dérivation verbale en berbère tachelhit*. Thèse de Doctorat, Université Paris 7
- Lahrouchi M (2010) On the internal structure of Tashlhiyt Berber triconsonantal roots. *Linguist Inq* 41(2):255–285
- Lahrouchi M (2013) A propos de I et de U en berbère: de la phonologie, de la morphologie et des phases. In: Tifrit A (ed) *Phonologie, Morphologie et Syntaxe*. Presses Universitaires de Rennes, Rennes, pp 21–30
- Lahrouchi M (2018) Syllable structure and vowel-zero alternations in Moroccan Arabic and Berber. In: Agwuele A, Bodomo A (eds) *The Routledge handbook of African linguistics*. Routledge, London, pp 168–180
- Lahrouchi M, Kern S (2018) From babbling to first words in Tashlhiyt language acquisition: a longitudinal two-case study. *Can J Linguist* 63(4):493–526
- Lahrouchi M, Ridouane R (2016) On diminutives and plurals in Moroccan Arabic. *Morphology* 26(3):453–475
- Lahrouchi M, Ségéral P (2010) Peripheral vowels in Tashlhiyt Berber are phonologically long: evidence from Tagnawt. *Brill Ann Afroasiat Lang Linguist* 2:202–212
- Lasri A (1991) *Aspects de la phonologie non-linéaire du parler berbère chleuh de Tidli*. Thèse de Doctorat, Université Paris III
- Lehiste I (1965) Juncture. In: Zwirner E, Bethge W (eds) *Proceedings of the fifth international congress of phonetic sciences*. Karger, Basel, pp 172–200
- Levi S (2011) Glides. In: van Oostendorp M, Ewen C, Hume E, Rice K (eds) *The Blackwell companion to phonology*, vol I. Wiley-Blackwell, Hoboken, pp 341–366
- Levin J (1985) *A metrical theory of syllabicity*. PhD dissertation, MIT
- Lowenstamm J (2008) On little n, √, and types of nouns. In: Hartmann J, Hegedüs V, van Riemsdijk H (eds) *Sounds of silence: empty elements in syntax and phonology*. Elsevier, Amsterdam, pp 105–144
- Marantz A (2001) *Words*. Ms. MIT, Cambridge
- Marantz A (2007) Phases and words. In: Choe S-H (ed) *Phases in the theory of grammar*. Dong In, Seoul, pp 191–222
- Marvin T (2002) *Topics in the stress and syntax of words*. PhD dissertation, MIT

- Marvin T (2013) Is word structure relevant for stress assignment. In: Matushansky O, Marantz A (eds) *Distributed morphology today: morphemes for Morris Halle*. The MIT Press, Cambridge (Mass), London, pp 79–93
- Mascaró J (1976) Catalan phonology and the phonological cycle. PhD dissertation, MIT
- Miller T (2018) Polysynthesis at the phonology-syntax Interface: a case study of Kiowa and Saulteaux Ojibwe. PhD dissertation, University of Delaware
- Mohanan PK (1982) Lexical phonology. PhD dissertation, MIT
- Newell H, Piggott G (2014) Interactions at the syntax-phonology interface: evidence from Ojibwe. *Lingua* 150:332–362
- Omari N (2012) Clitic placement in Amazigh. *Asinag* 7:185–200
- Ouali H (2005) Verb movement and object pronominal clitic distribution in Berber. Ms., The University of Michigan
- Ouali H (2011) Agreement, pronominal clitics and negation in Tamazight Berber. Continuum International Publishing Group, London/New York
- Ouhalla J (1988) The syntax of head movement: a study of Berber. PhD dissertation, University College London
- Padgett J (2008) Glides, vowels and features. *Lingua* 118:1937–1955
- Pater J (2010) Morpheme-specific phonology: constraint indexation and inconsistency resolution. In: Parker S (ed) *Phonological argumentation: essays on evidence and motivation*. Equinox, Sheffield, pp 123–154
- Pike K, Pike E (1947) Immediate constituents of Mazateco syllables. *Int J Am Linguist* 13:78–91
- Puech G, Louali N (1999) Syllabification in Berber: the case of Tashlhiyt. In: Ohala J (ed) *Proceedings of the 14th international congress of phonetic sciences*. University of California Press, San Francisco, pp 747–750
- Ridouane R (2003) Suites de consonnes en berbère: phonétique et phonologie. Thèse de Doctorat, Université Paris 3-Sorbonne Nouvelle
- Ridouane R (2008) Syllables without vowels: phonetic and phonological evidence from Tashlhiyt Berber. *Phonology* 25:321–359
- Ridouane R (2009) Pharyngealization and pharyngeal Tashlhiyt in Berber: a descriptive and comparative study. *International Workshop on Pharyngeal and Pharyngealization*, Newcastle
- Ridouane R (2014) Tashlhiyt Berber. *J Int Phon Assoc* 44:207–221
- Rosenthal S (1994) Vowel-glide alternation in a theory of constraint interaction. PhD dissertation, University of Massachusetts
- Samuels B (2010) Phonological derivation by phase: evidence from Basque. *Univ Penn Work Pap Linguist* 16:166–175
- Scheer T (2008) Why the prosodic hierarchy is a diacritic and why the interface must be direct. In: Hartmann J, Hegedüs V, van Riemsdijk H (eds) *Sound of silence: empty elements in syntax and phonology*. Elsevier, Amsterdam, pp 146–192
- Scheer T (2011) A guide to morphosyntax-phonology interface theory. Walter de Gruyter, Berlin
- Scheer T (2015) Précis de structure syllabique. ENS Editions, Lyon
- Selkirk E (1982) Syllables. In: van der Hulst H, Smith N (eds) *The structure of phonological representations*. Foris, Dordrecht, pp 337–383
- Selkirk E (1984) On the major class features and syllable theory. In: Aronoff M, Oehrle RT (eds) *Language sound structures*. MIT Press, Cambridge, MA, pp 107–136
- Soutsane K (2008) Syllabification and phonological rule application in Tashlhiyt Berber. PhD dissertation, Durham University
- Trager G (1962) Some thoughts on juncture. *Stud Linguist* 16(1):11–22
- Yeou M (1997) Locus equations and the degree of coarticulation of Arabic consonants. *Phonetica* 54(3–4):187–202
- Zeroual C (2000) *Propos controversés sur la phonétique et la phonologie de l'arabe marocain*. Thèse de doctorat, Université Paris 8

The Phonetics and Phonology of Tashlhiyt Geminates: An Overview

5

Rachid Ridouane

Contents

Introduction	104
Representation of Consonant Gemination	105
Singletons Versus Geminates: Acoustic and Articulatory Attributes	108
Primary Correlate of Gemination	108
Secondary Correlates of Gemination	110
A Contrast with Multiple Correlates	111
The Perception of Geminates in Tashlhiyt	112
Additional Issues	114
Lexical Versus Derived Geminates	114
Gemination and Speech Rate: A Search for Invariance	115
Conclusion	116
References	117

Abstract

The length of consonants plays an important role in distinguishing words in various languages around the world, by contrasting singletons and geminates. A growing body of research has been conducted on the phonetics and phonology of this contrast over the last few decades. This chapter offers a general overview into this research, focusing on gemination in Tashlhiyt. It summarizes previous and ongoing studies, discussing the representation of gemination in the phonological system, as well as how the physical characteristics of geminates contribute to their unique patterns. Tashlhiyt is particularly interesting as it has contrastive geminates in word-initial, medial, and final positions, and displays various types of geminates such as lexical, morphological, and phonological.

R. Ridouane (✉)

Laboratoire de Phonétique et Phonologie, CNRS & Sorbonne Nouvelle, Paris, France

e-mail: rachid.ridouane@sorbonne-nouvelle.fr

Keywords

Gemination · Length contrast · Duration · Tenseness · Autosegmental phonology · Acoustics · Articulation · Perception

Introduction

Tashlhiyt features a lexical contrast between singletons and geminates for all consonants. This contrast is attested in various positions within words, as shown in the examples in (1) (Geminates are not infrequent in Tashlhiyt words, although still far less frequent than their singletons counterparts. In Alderete et al. (this volume), based on a corpus of nine texts with 18,827-word tokens, 19% out of 37,334 sound tokens are geminates. Some geminate consonants are, however, marginally attested in the lexicon, especially, pharyngeals, laryngeals, labialized uvulars, and pharyngealized coronals.).

(1)	[tut]	“she hit”	[ttut]	“forget him”
	[tidi]	“sweat”	[tiddi]	“height”
	[ifis]	“hyena”	[ifiss]	“he is silent”
	[kkstt]	“take it (fem) off”	[kst]	“feed it (m) on”

In Tashlhiyt, geminates are not only found in intervocalic position, but also in initial and final positions. This feature is relatively rare cross-linguistically as noted in studies by Taylor (1985), Thurgood (1993), Muller (2001), and Dmitrieva (2012). Even rarer, geminates may occur with one or more consonants before or after them (e.g., [tkkstntt] “you took them off”), and words can consist only of a geminate (e.g., [ʃʃ] “eat” and [ggʷ] “wash”).

In addition to the geminates that are present in the lexicon, Tashlhiyt also has geminates that are derived phonologically or morphologically. Phonological geminates can be formed through total assimilation between two adjacent segments, such as between /d/ and /k/ in /rad-k awi-k/ [rakk awiʔ] “I will take you,” or from the concatenation of two identical segments across a word boundary, for example in /tuf = fas/ [tuffas] “it’s better than Fès.” Morphological geminates are formed through processes such as pluralization (2a) or the formation of imperfective forms (2b) (On imperfective gemination, see Dell and Elmedlaoui (1991, 2013), Jebbour (1999), Bensoukas (2001), Lahrouchi (2008, 2010), among others.).

(2)	a.	<i>Singular</i>	<i>Plural</i>	
		afus	ifassn	“hand”
		adʕar	idʕarn	“foot”
		afud	ifaddn	“knee”
	b.	<i>Perfective</i>	<i>Imperfective</i>	
		gru	grru	“gather”
		nkr	nkkr	“stand up”
		ftu	fttu	“go”

This chapter focuses on the phonetics and phonology of gemination. The term “gemination” will be used in a generic sense with a distinction between phonological length and tension. Phonetic data will be evaluated in this sense, with phonetic duration referring to phonological length, and other correlates (e.g., intensity of the release) referring to the [tense] feature. A central contentious issue is the implementation of the gemination contrast in Tashlhiyt through duration, despite the high variability of this attribute. Additionally, this chapter examines how other correlates may enhance the contrast when duration is not clearly perceptible.

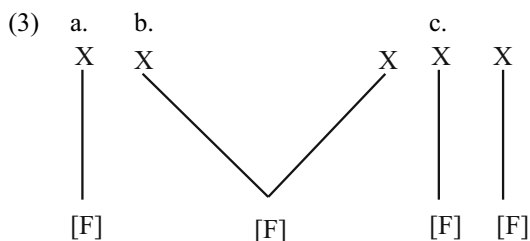
The structure of this chapter is as follows: section “[Representation of Consonant Gemination](#)” discusses the nature of the phonological representation of gemination, including an overview of arguments provided by proponents of a sequential approach, and proponents of a feature distinction. Section “[Singletons Versus Geminates: Acoustic and Articulatory Attributes](#)” discusses the acoustic and articulatory correlates of gemination, and the necessity for positing one primary attribute – duration – and additional secondary attributes. Section “[The Perception of Geminates in Tashlhiyt](#)” discusses the perception of geminates, with a special focus on the contrast between singleton and geminate voiceless stops in word-initial position. Section “[Additional Issues](#)” provides a summary of two additional issues not dealt with in detail in this chapter: the phonetic and phonological patterns of different types of geminates, and the effect of speech rate in the way the singleton/geminate contrast is acoustically realized. Section “[Conclusion](#)” concludes this chapter with a brief summary and an outlook on areas that require further research.

Representation of Consonant Gemination

One of the foundational questions in geminate phonology is whether a geminate is a single unit or a sequence of two identical units. This question, first posed in the 1930s (e.g., Swadesh 1937) and widely discussed since then, relates to the ambiguous nature of geminates: they sometimes act as a single consonant and sometimes as a sequence of two consonants. In “Sound Pattern of English,” Chomsky and Halle (1968) proposed two ways of phonologically representing geminates: as a segment with the [+long] feature or as a sequence of two segments with identical feature bundles. However, subsequent work has highlighted the limitations of these descriptions and have shown that neither representation fully captures the behavior of geminates in relation to certain phonological processes (Kenstowicz 1970; Pyle 1971; Kenstowicz and Pyle 1973).

The previous work done within the framework of standard Generative Phonology did not provide a clear answer to the problem of representing geminates, but it did succeed in identifying this issue and in proposing a generalization that would prove to be quite promising (Kenstowicz 1994). Kenstowicz (1970), for instance, observed that geminates are typically treated as a single unit by quality-sensitive rules (rules that are sensitive to the internal makeup of segments) and as two units by quantity-sensitive rules (rules that are sensitive to the number of segments). This distinction

between quantitative and qualitative rules is further developed in the framework of CV Phonology (Leben 1980; Clements and Keyser 1983). The central idea in CV phonology is that syllable positions are represented on a prosodic tier (typically marked as X slots), separate from the melodic tier (containing segmental features [F]). The two tiers are connected by association lines. This approach allows for the following structural representations for geminates in Tashlhiyt (Dell and Elmedlaoui 1997; Ridouane 2003) (The moraic representation is a different way of structuring geminates, where it is assumed that a geminate has an inherent moraic structure. Jebbour (1996) provides the moraic representation of Tashlhiyt geminates.):



Representation (3a) illustrates a singleton consonant, (3b) illustrates a geminate consonant, and (3c) illustrates a sequence of two adjacent consonants. Geminates are differentiated from singletons not by a specific feature, but by the number of prosodic positions they occupy: the singleton is linked to one X position (3a), while the geminate is linked to two X positions. Representation (3b) effectively captures the ambiguity of geminates: it is similar, on one hand, to that of singleton consonants as both have a single melodic position (compare 3b to 3a), and on the other hand, to that of a sequence of two adjacent consonants as both have two X positions (compare 3b to 3c).

Another significant question related to the representation of geminates concerns the nature of the phonetic contrast between singletons and geminate: is it based on duration or tenseness? Much of the research within Autosegmental Phonology has not delved deep into this question (but see Clements 1986). Indeed, marking the distinction between one or two prosodic positions at the prosodic level is sufficient to account for the differing behavior of singletons and geminates in regard to phonological processes. On the other hand, those who argue that geminates and singletons are distinguished by a distinctive feature are naturally interested in determining the nature of this feature (i.e., whether the feature is [+/- long] or [+/- tense]).

The feature [tense] has been used in a variety of ways in relation to different consonantal systems (see Jessen (1998), for a review). Linguists who use this feature to define gemination in Amazigh typically use it to indicate increased strength or articulatory energy (Mitchell 1957; Applegate 1958; Galand 1953, 1997; Chaker 1975; Ouakrim 1994; Louali and Puech 1994) (See also Jakobson, Fant and Halle (1952), and Jessen (1998), who define geminates by the feature [tense]; and Kohler (1984) who proposes the feature [fortis] to characterize this opposition.). Louali and Puech (1994) conducted acoustic, perception, and aerodynamic analyses to

determine whether the geminate stops in Tashlhiyt are long or tense segments. They found that the duration of geminates is always longer than that of the corresponding singleton consonants. The perceptual study revealed that native speakers primarily use duration as a cue to distinguish geminates from singletons, but they also use additional cues in a secondary manner. The aerodynamic study highlighted duration as a distinctive parameter, and also found that geminates exhibit a fast and significant rise in oral pressure at the time of release, and that the energy of the release is more intense for geminates.

Ouakrim (1994) makes a clear distinction between geminates, which he believes should only refer to sequences of two identical consonants separated by a morphological boundary (i.e., representation (3c) above), and tense consonants, which cannot be broken down into two phonetic segments or between two syllables. According to him, the longer duration of tense segments is a result of the increased physiological energy required for their production, whereas during the holding of geminates, the speaker subjectively maintains the same effort as in non-geminates or non-tense consonants. One main characteristic of tenseness is the shortening of the preceding vowels, which Ouakrim (1994) did not observe in the context of heteromorphemic geminates.

Galand (1953) proposed that geminates are characterized by muscle tension, and presented a set of arguments in support of this idea in a later work (Galand 1997). One key observation is that geminates occur in positions (e.g., initial and final) where it is not possible to interpret them as belonging to two adjacent syllables (e.g., [kks] “take off” [juʃf] “it is swollen”). Galand also points out that muscle tension appears to be the most likely feature to explain patterns of geminate distribution observed in Tashlhiyt and other Amazigh varieties. For example, when a singleton and its geminate counterpart have different [continuant] values, it is always the geminate that is [– continuant] and the singleton that is [+ continuant] (e.g., /x/ and /kk/ in the spirantizing Amazigh varieties). Additionally, in cases where a singleton and its geminate counterpart have realizations that differ in voicing, it is always the singleton that is [+ voiced] and the geminate that is [– voiced] (e.g., ([dʰ] ≈ [tʰ] in [dʰr] “fall, *perfective*” and [tʰar] “fall, *imperfective*”). Galand posits that duration alone cannot explain these phenomena and suggests that mismatches in muscle tension may play a role. He argues that this muscle tension can inhibit vocal cord vibration and account for the perceived forcefulness of geminates. He uses examples of minimal pairs such as [krz] “plow, *perfective*” versus [kkrz] “plow, *imperfective*” to support his view. In this case, he suggests that it would be difficult to differentiate the two forms based on duration alone as the two contrastive voiceless stops /k/ and /kk/ are in the initial position and nothing is heard before their release. Thus, he argues that only a variation in muscle tension can account for more forceful explosion of /kk/ compared to /k/. Lastly, Galand emphasizes that muscle tension can occur in all positions, whereas duration only applies in certain positions.

Adopting an autosegmental approach, Dell and Elmedlaoui (1997, 2002) argue that the difference between singletons and geminates is structural, and that geminates are a single bundle of distinctive features linked to two prosodic positions (as in (3b) above). They present several arguments in favor of this analysis. One argument is that the

fusion of two identical singleton consonants into one geminate consonant accounts for the homophony between, for example, /gn-n/ “they slept” and /g = nn/ “put there.” Another argument is the fission of some final geminates in the imperfective (e.g., when the perfective [bdd] “to stand up” yields [ttbdad] in the imperfective). They have also observed that geminates and sequences of two consonants tend to behave similarly in certain phonological processes, as previously noted by researchers such as Guerssel (1977) and Saib (1977). They also provide additional evidence from syllable structure and morphology (Dell and Elmedlaoui 2002: 41–55).

Several experimental studies have been conducted to test the representational model of geminate consonants within Autosegmental Phonology (Ridouane 2007, 2010; Ridouane and Hallé 2017). The findings of these studies showed that the model accurately explains a range of observations: (i) the distinction between singletons and geminates is primarily a temporal distinction, including for voiceless stops in utterance-initial position, (ii) when this temporal difference is not present in the acoustic signal, native listeners have more difficulty in distinguishing between singletons and geminates, and (iii) the comparison of different types of geminates (lexical, derived through assimilation, and derived through concatenation) shows that only true geminates exhibit phonetic characteristics that are consistent with the representational model. The following sections elaborate on these three observations.

Singletons Versus Geminates: Acoustic and Articulatory Attributes

The contrast between singleton and geminate consonants in different languages of the world has been the subject of multiple studies, particularly in intervocalic position (see Ridouane (2007), for a review of 24 languages, see also Kawahara (2015), Hamzah et al. (2016), and Kubozono (2017)). A consistent finding among these studies is that geminates are systematically longer than their singleton counterparts. In addition to consonant duration, gemination may also have implications for other phonetic aspects of a form such as burst amplitude, quality and duration of adjacent vowels, among others. The following section presents a review of the phonetic characteristics of geminates in Tashlhiyt, drawing on both acoustic and articulatory data, and distinguishing primary from secondary phonetic attributes of gemination.

Primary Correlate of Gemination

The primary correlate of gemination in Tashlhiyt is duration, meaning that geminates are consistently longer than their singleton counterparts, regardless of segment type and position in the word (Ridouane 2007). Figure 5.1 shows acoustic waveforms and spectrograms of the pair /tidi/ “sweat” and /tiddi/ “height” contrasting /d/ to /dd/ in intervocalic position. As the figure shows, the closure duration of geminate /dd/ is more than twice as long as the closure phase of singleton /d/. Figure 5.2 illustrates

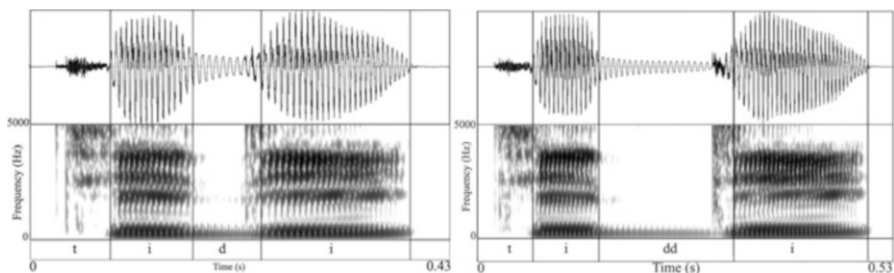


Fig. 5.1 Acoustic waveform and spectrogram of the forms [tidi] “sweat” (left) and [tiddi] “height” (right), illustrating the durational differences between intervocalic [d] and [dd]

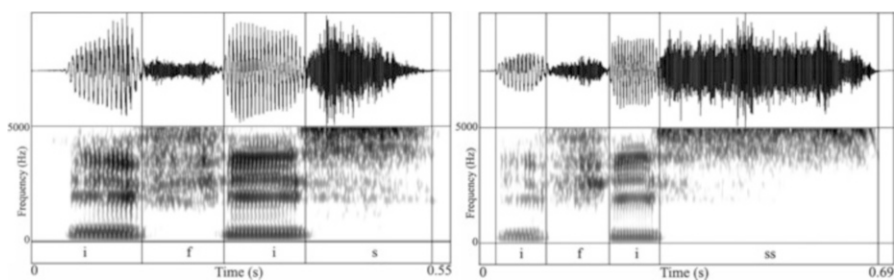


Fig. 5.2 Acoustic waveforms and spectrograms of the forms [ifis] “hyena” (left) and [ifiss] “he is silent” (right), illustrating the durational differences between word-final [s] and [ss]

similar duration differences in word-final position for the voiceless fricatives /s/ and /ss/ in [ifis] “hyena” and [ifiss] “he is silent.”

The size of duration differences between geminates and singletons in Tashlhiyt vary depending on consonant type and position within the word. For instance, geminate stops are more than twice as long as their singleton counterparts, while geminate fricatives are slightly less than twice as long. Additionally, the difference between geminates and singletons is greater in word-initial and word-final positions compared to word-medial position, resulting in a higher geminate-singleton ratio in non-medial positions.

As the closure duration for voiceless stops cannot acoustically distinguish between singletons and geminates when they are at the beginning of an utterance (e.g., [tut] “she hit” vs [ttut] “forget it”), it is worth considering if speakers maintain the contrast at the articulatory level. Ridouane (2007) conducted research using electropalatographic (EPG) data and found that speakers do indeed maintain the duration differences between singletons and geminates, even for voiceless stops in initial position (as illustrated in Fig. 5.3).

The EPG data presented in Ridouane (2007) dealt with the temporal dimension of the singleton/geminate distinction. More recently, Ridouane and Hallé (2017) showed that these differences in contact duration could be accompanied by differences in contact area: voiceless and voiced stops, but not fricatives, are produced

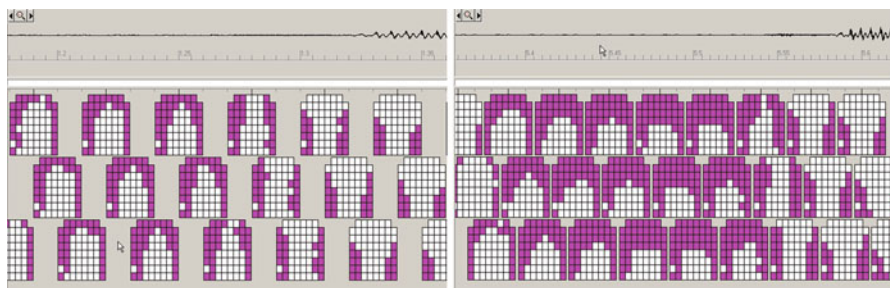


Fig. 5.3 An illustration of durational and spatial differences between utterance-initial singleton /t/ and geminate /tt/ using EPG data. The figure shows the electrodes activated during the closure phase for the two stops in [tɪli] “ewe” (left) and [ttɪli] “have” (right)

with larger contact area for geminates than for singletons (see also Fig. 5.3). The observed differences for stops are probably an automatic consequence of the longer contact duration for geminates compared to singletons. This suggests that geminate stops would not differ from singletons in terms of the underlying gestural target, but have more time to reach this target. The lack of spatial differences between simple and geminate fricatives is consistent with this analysis. Differences in contact duration for dental fricatives cannot result in contact amplitude differences similar to those for stops because the articulation of these fricatives is more strictly controlled: dental fricatives cannot be produced with a different tongue-palate contact area without risking confusion with the alveopalatal fricatives. In other words, speakers control tongue-tip location more strictly when producing /s/, so that /ʃ/ is not produced instead. This is consistent with what has been reported for singletons and geminates in Italian (Payne 2006) and in Cypriot Greek (Armosti 2009).

Secondary Correlates of Gemination

In addition to consonant duration, Tashlhiyt uses other cues to distinguish between singletons and geminates. These cues can be either temporal or non-temporal. One notable temporal cue is the shortening of the duration of the vowel preceding the geminate consonant (see Figs. 5.1 and 5.2). This cue is considered secondary as it is only implemented when the geminate is actually preceded by a vowel and cannot be used in initial position or in words consisting of only a geminate (e.g., [kk] “cross”). Vowel shortening before geminates is observed in other languages like Moroccan Arabic (Zeroual 2006), Bengali (Lahiri and Hankamer 1988), Italian (Esposito and Di Benedetto 1999), Hindi (Ohala 2007), and Malayalam (Local and Simpson 1999), and is often attributed to differences in syllabic structure between singletons and geminates. But Ridouane (2010) study showed that the vowel remains shorter in geminate context compared to non-geminate VCCV context, although they share the same syllable structure. Additionally, the syllabic explanation is unable to account for the shortening observed in word-final position, such as in the case of the vowel /i/

in [ifis] compared to [ifiss] shown in Fig. 5.2. The vowel [i] is in a closed syllable in both the singleton and geminate context, yet it is only shortened in the geminate context.

Another interpretation for this shortening was put forward by Malécot (1968, 1970) and adapted by Ouakrim (1994) for Tashlhiyt. According to Malécot (1970), vowel shortening is due to the tendency of speakers to anticipate an important effort and to delay the less important one, so that the more energy a consonant requires, the shorter the preceding vowel. The shortening of the vowel would, thus, be an indication of the greater energy required to produce a geminate. Related to this interpretation, and probably the one most widely held, is that which considers that the length contrast on a geminate segment is enhanced by making the preceding segment shorter (Kluender et al. 1988).

The duration of stop release is another temporal parameter that can implement the singleton/geminate contrast. Ridouane (2007) has shown that Tashlhiyt voiced geminate stops exhibit a longer release duration, on average 30 ms, compared to their singleton counterparts (17 ms on average). This phenomenon has been observed in multiple languages, including Bengali (Mikuteit and Reetz 2007) and Finnish (Doty et al. 2007; Engstrand and Krull 1994), and is likely attributed to the devoicing that occurs during the production of voiced geminates. In contrast, no significant difference in voice onset time (VOT) duration was found for voiceless stops. This can be explained by the fact that singleton and geminate voiceless stops are produced with a similar degree of glottal opening at the time of oral release, as established by previous research (Ridouane et al. 2006).

Gemination can have a range of effects on nontemporal parameters of speech sounds. These include, but are not limited to, the amplitude of release, the nature of the closure, and the degree of lenition. Research has shown that geminate stops tend to be produced with a higher amplitude of release than singletons (Louali and Puech 1994; Ridouane 2007). However, this tendency is speaker-dependent, with some speakers exhibiting the effect, while others do not. Similarly, singleton stops may be produced with an incomplete closure, while geminate stops are systematically produced with total occlusion. The degree of lenition during the holding phase of singleton stops varies depending on the voiced/voiceless nature and place of articulation of the consonant. For instance, voiced velar stops are more likely to occur without full closure compared to other consonants. Additionally, partial devoicing can occur in voiced geminates, but the extent of this effect varies across speakers, places of articulation (with velar stops being more likely to exhibit partial devoicing), and context (more common in final position).

A Contrast with Multiple Correlates

In summary, the phonetic implementation of gemination contrast in Tashlhiyt can be characterized by various correlates. The primary correlate is consonant duration, as it is attested in all contexts examined, even for voiceless stops in utterance-initial position. There is no evidence that the longer duration of geminates is a result of tense

articulation, rather it appears that the differences in duration are the result of speakers' intention to maintain a longer duration for geminates (see Louali and Maddieson 1999). The observed differences in the duration of release and the degree of devoicing can be considered as concomitant correlates, being consequences of the devoicing that occurs as a result of the longer duration of these segments. Vowel shortening and the amplitude of release, which may be interpreted as manifestations of tense articulation, are secondary correlates. They are either contextually limited (vowel shortening) or variable across speakers (amplitude of release). These secondary correlates can be considered as enhancing attributes of gemination (Keyser and Stevens 2006; Stevens and Keyser 2010; Clements and Ridouane 2006), as they reinforce the primary correlate by adding additional acoustic properties that increase the perceptual distance between the two phonemic categories. These enhancing correlates can be exploited in cases where the primary correlate is not perceptually recoverable, such as voiceless stops after a pause, where listeners cannot detect the differences in closure duration between singletons and geminates (see section “[The Perception of Geminates in Tashlhiyt](#)”).

In Tashlhiyt, the durational differences between singletons and geminates can be accurately represented by structurally treating geminates as two units of duration linked to one melodic position, as shown in (3b) above. However, the representation of the enhancing features of gemination remains unclear. One possibility is to assign the feature [tense] to the geminate representation as a general effect of phonetic implementation rules. From an articulatory phonology perspective (Browman and Goldstein 1992), this feature may not be necessary as the singleton can be viewed as an undershot version of the geminate, with both having the same target specifications but the singleton not reaching the target value due to its shorter duration. An alternative interpretation is suggested by Pierrehumbert's (2002) hybrid exemplar model, which posits that memory traces of geminates are hyperarticulated, as only hyperarticulated examples are reliably recognized in competition with minimally different competitors. This predicts differential effects on additional correlates depending on whether the geminate is in a dense lexical neighborhood with singleton competitors or not (Pierrehumbert and Clopper 2010).

The Perception of Geminates in Tashlhiyt

The perception of gemination contrast in Tashlhiyt has been the subject of limited research. Previous studies by Louali and Puech (1994), Ouakrim (1999), and Ridouane and Hallé (2017) have investigated the cues used to perceive this contrast. In intervocalic position, duration was found to be the primary cue for perception, with other parameters such as preceding vowel duration and stop release amplitude serving as secondary cues (Louali and Puech 1994). However, in initial position, the gemination contrast raises questions about whether durational differences can be perceived for voiceless stops. Ouakrim (1999) conducted a perceptual study

examining the minimal pair [tutas] “she hit for him” versus [tttutas] “you forgot for him,” manipulating the release portions (corresponding to VOT) of singleton /t/ and geminate /tt/. He also cross-spliced the vowels following the singleton and geminate stops to determine whether VOT and, or, the following vowel play a significant role in the perception of gemination contrast in this position. His results showed that listeners tended to cross the original meanings of the pair, suggesting that they can reliably perceive the singleton/geminate contrast in the absence of acoustic duration differences.

Ridouane and Hallé (2017) conducted an AXB discrimination experiment to investigate the perception of gemination contrasts in Tashlhiyt. The study included comparisons of three types of word-initial consonants: voiceless stops (e.g., [tut] “she hit” vs. [ttut] “forget him”), voiced stops (e.g., [gar] “bad” vs. [ggar] “be last”), and voiceless fricatives (e.g., [fit] “give it” vs. [ffit] “pour it”). The study predicted that native listeners could distinguish [gar] from [ggar] as the duration of voicing murmur serves as a reliable acoustic cue, and could distinguish [fit]–[ffit] based on differences in friction noise duration. However, for contrasts such as [tut]–[ttut], listeners may give conflicting and unreliable perceptual judgments, given the absence of robust acoustic cues for gemination, particularly closure duration. The study included 23 native speakers/listeners and results showed that they performed near ceiling level on word-initial fricatives and voiced stops (>95%). However, for word-initial voiceless stops, the listeners performed less than 62% correct discrimination. These results were corroborated by reaction time data, which also showed longer RTs for voiceless stops than voiced stops, or fricative contrasts by some 180 ms (see Fig. 5.4).

The results of the study conducted by Ridouane and Hallé (2017) indicate that listeners have difficulty correctly discriminating a voiceless singleton stop from its geminate counterpart in the absolute initial position, suggesting that a phonological contrast cannot be perceived in the absence of clear acoustic consequences. The study’s findings have implications for the role of duration as the primary correlate of the opposition between singletons and geminates. Specifically, duration is primary

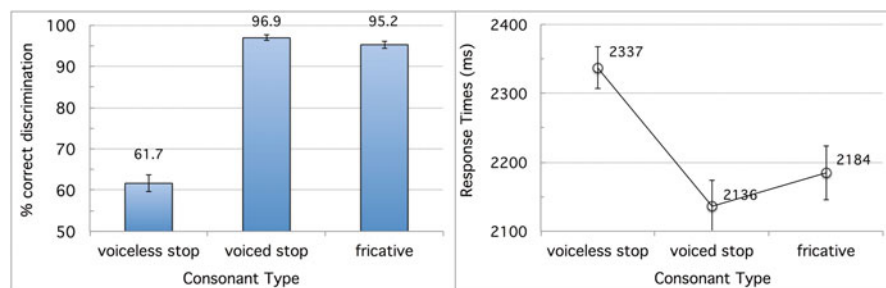


Fig. 5.4 Correct discrimination rate (left) and reaction time (right) for word-initial singleton-geminate contrasts for the three onset consonant types. Error bars represent standard errors. (Adapted from Ridouane and Hallé 2017)

because it is consistently present in all contexts where the contrast occurs, and it serves as the most important perceptual cue for native listeners to accurately perceive the contrast at a native listening level.

Additional Issues

The phenomenon of gemination raises other issues that have not been explored in this chapter. In this last section, I will briefly address two issues: the distinctions among different types of geminates and the effect of speech on the durational attributes of gemination.

Lexical Versus Derived Geminates

As previously mentioned, the surface geminates in Tashlhiyt can originate from various sources. Tautomorphemic lexical geminates are found in the lexicon and represented at the underlying level as a single melodic unit linked to two prosodic positions (3b). Heteromorphemic geminates can result from complete assimilation or from the concatenation of two identical consonants separated by a word boundary. Geminates that result from total assimilation are also represented as a single melodic unit linked to two prosodic positions, due to autosegmental propagation (Hayes 1986a). Geminates that result from concatenation are represented at the underlying level as a sequence of two prosodic units, each linked to a melodic position. These “fake” geminates, according to McCarthy (1986), are identical to lexical geminates at the surface level due to tier conflation. If these analyses are correct, all three types of geminates, regardless of their underlying representations, should be identical at the surface level and represented as a single melodic unit linked to two prosodic positions.

Ridouane (2010) investigated these issues using acoustic data and found that the three types of geminates exhibit similar consonantal durations. This supports the idea of an identical representation for these three types of geminates at the prosodic tier, which is associated with two timing units. A lack of distinctions in consonant duration between true and fake geminates was found to be consistent across several languages, including Estonian (Lehiste, et al. 1973), Levantine Arabic (Miller 1987), and Bengali (Lahiri and Hankamer 1988). Despite the indistinguishable absolute consonant duration, it was found that vowels are significantly shorter before lexical geminates and assimilated geminates than before concatenated geminates. Additionally, measurements revealed a tendency for lower amplitude of release for concatenated geminates compared to lexical and assimilated geminates.

The fact that geminates resulting from total assimilation are categorically identical to lexical geminates provides further evidence that the external sandhi process of total assimilation is correctly represented within the framework of a non-linear model, which expresses assimilation as an autosegmental propagation, rather than as a modification of a segment (see Ladd, et al. (2003) for similar results in Sardinian). This model also provides an adequate representation for concatenated

gemimates, assuming they are represented at the surface level as two timing positions, each linked to a melodic position.

Obviously, the mere fact of having a sequence of two identical consonants is not sufficient in itself to manifest the expected characteristics of a “true” geminate. The relationship between the phonetic characteristics of different types of gemimates and their phonological behavior is demonstrated by the failure of certain phonological processes to alter true gemimates while affecting fake gemimates (Hayes 1986b; Elmedlaoui 1993). In Tashlhiyt, spoken in Western High Atlas, for example, lexical gemimates and gemimates resulting from total assimilation never undergo spirantization, while gemimates resulting from concatenation do spirantize (Ridouane 2010).

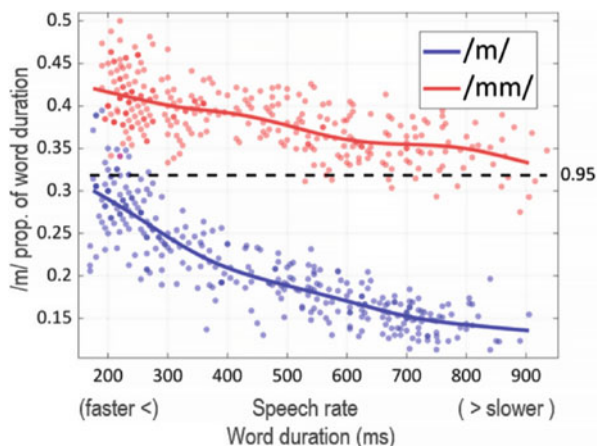
Gemination and Speech Rate: A Search for Invariance

Several studies have demonstrated that speech rate affects singletons and gemimates differently (Pind 1995; Pickett et al. 1999; Hirata and Whiton 2005). The research question was whether there exists an invariant acoustic correlate that can transcend these speech rates. The findings suggest that this invariant measure is relational, rather than absolute. Pickett et al. (1999) found that the ratio of consonant duration to preceding vowel duration (C/V) in Italian can discriminate between singletons and gemimates, both within and across different speaking rates. Hirata and Whiton (2005) found similar results in Japanese, with the highest accuracy for the consonant-to-word ratio (C/W) (see also Idemaru and Guion 2008).

A recent preliminary study by Hermes et al. (2021) aimed to investigate the feasibility of identifying an invariant acoustic attribute of gemination contrast in Tashlhiyt that is independent of speech rate. The study employed a cross-linguistic comparative approach, using data from four typologically unrelated languages: Finnish, Italian, Japanese, and Tashlhiyt. To elicit variation in speech rate, a motion-based cue, in the form of a red box moving across the screen, was used. The cue duration ranged from a minimum of 0.75 s to a maximum of 3 s in 20 steps. The study included one speaker per language, who was recorded producing the pair /ima/ and /imma/ embedded in a carrier sentence, repeated 640 times. The acoustic parameters measured included both absolute and relational measures, such as preceding vowel duration, consonant duration, following vowel duration, C/V ratio, and C/W ratio.

The results of the study showed that the durations of [m] and [mm] in Tashlhiyt, as well as in the other languages included in the study, change with rate, with geminate consonants increasing more as rate slows. When speech rate is slower, the differences in absolute durations between singletons and gemimates are further enhanced with an increase in geminate durations. The rate effects on the preceding vowel were found to be smaller before [mm] than [m], while the interaction with the following vowel was similar between singletons and gemimates. Importantly, despite the large overlap of singleton/geminate durations at fast rates, a relational measure – the C/W ratio – reliably distinguished between the two categories and could form the basis for a speech rate-independent attribute of gemination. As shown in Fig. 5.5, the C/W ratio allowed for accurate classification of singleton and geminate tokens in

Fig. 5.5 Ratio of /m/ and /mm/ to word duration plotted against word duration in Tashlhiyt. Dashed line represents optimal boundary computed as 0.32 of the word duration (0.95 indicates the classification accuracy)



Tashlhiyt, with a boundary at 32% and a classification accuracy of 95%. This means that if the duration of the labial nasal is less than 32% of the word duration, it is almost certainly a singleton /m/, and if it is more than 32% of the word duration, it is almost certainly a geminate /mm/. Interestingly, the boundary ratio was found to be similar across the four languages (varying from 29% to 37%). This provides support for previous work, and suggests that a cross-linguistic rate-independent attribute of gemination is not absolute but relational.

Conclusion

The present chapter provides an overview of the nature of gemination contrast and its phonetic implementation and perception in Tashlhiyt. From a phonological perspective, geminates in this language are represented within the framework of Autosegmental theory, as sequences of two timing positions linked to a single melodic position. Empirical data from both production and perception studies allowed to test a set of predictions that follow from this phonological representation. It was shown that the distinction between singletons and geminates is primarily a temporal one, and this holds true for all consonants in all positions, including voiceless stops after a pause. Additionally, it was found that in the absence of this temporal dimension, the distinction between singletons and geminates is perceptually much weaker, with native listeners unable to adequately recover the contrast for post-pausal voiceless stops. The autosegmental analysis also provided an adequate representation for different types of geminates and accounts for the fact that geminates resulting from total assimilation are categorically identical to lexical geminates.

The review of the current state of research on gemination in Tashlhiyt shows that this issue has been examined from both theoretical and empirical perspectives, however, further research is still needed to fully understand the behavior of these

segments. Some potential areas for future study include the acquisition of gemination contrast by native speakers, the dynamic articulatory characteristics of geminates, the syllabic status of initial and final geminates, and the processing of consonant length contrast.

References

- Applegate JR (1958) An outline of the structure of Shilha. American Council of Learned Societies, New York
- Armosti S (2009) The Phonetics of plosive and affricate gemination in Cypriot Greek. PhD thesis, University of Cambridge
- Bensoukas K (2001) Stem forms in the nontemplatic morphology of Berber. Dissertation, Mohammed V University, Rabat
- Browman CP, Goldstein L (1992) Articulatory phonology: an overview. *Phonetica* 49(3–4):155–180
- Chaker S (1975) Les paramètres acoustiques de la tension consonantique en berbère (kabyle). *Trav Inst Phonétique Aix* 2:151–168
- Chomsky N, Halle M (1968) The sound pattern of English. Harper and Row, New York
- Clements GN (1986) Compensatory lengthening and consonant gemination in LuGanda. In: Wetzels L, Sezer E (eds) *Studies in compensatory lengthening*. Foris, Dordrecht, pp 37–77
- Clements GN, Keyser SJ (1983) *CV phonology: a generative theory of the syllable*. MIT Press, Cambridge
- Clements GN, Ridouane R (2006) Part I: quantal phonetics and distinctive features: a review. Part II: distinctive feature enhancement: a review. In: *Proceedings of ISCA tutorial and research workshop on experimental linguistics*, pp 17–24 et 28–30
- Dell F, Elmedlaoui M (1991) Clitic ordering, morphology and phonology in the verbal complex of Imdlawn Tashlhiyt Berber. *Lang Orient Anc Philol Linguist* 3:77–104
- Dell F, Elmedlaoui M (1997) Les géménées en berbère. *Linguist Afr* 19:5–55
- Dell F, Elmedlaoui M (2002) Syllables in Tashlhiyt Berber and in Moroccan Arabic. Kluwer, Dordrecht
- Dell F, Elmedlaoui M (2013) Syllables and gemination in imperfective stems in Tashlhiyt Berber. *Brill J Afr Lang Linguist* 5(1):1–34
- Dmitrieva O (2012) *Geminate typology and the perception of consonant duration*. Doctoral dissertation, Stanford University
- Doty CS, Idemaru K, Guion SG (2007) Singleton and geminate stops in Finnish- Acoustic correlates. In: *Proceedings of the 8th Annual Conference of the International Speech Communication Association*. Curran Associates, Antwerp, pp 2737–2740
- Elmedlaoui M (1993) Gemination and spirantization in Hebrew, Berber, and Tigrinya: a fortis-Lenis module analysis. *Linguist Commun* 1:121–176
- Engstrand O, Krull D (1994) Durational correlates of quantity in Swedish, Finnish and Estonian: data from spontaneous speech. Working papers 43. Lund University, Department of Linguistics, Lund, pp 54–57
- Esposito A, Di Benedetto M (1999) Acoustical and perceptual study of gemination in Italian stops. *J Acoust Soc Am* 106:2051–2062
- Galand L (1953) La phonétique en dialectologie berbère. *Orbis* II(1):225–233
- Galand L (1997) Les consonnes tendues du berbère et leur notation. In: Taifi M (ed) *Voisinage, Mélanges en hommage à la mémoire de Kaddour Cadi*. Publications de la Faculté des Lettres et des Sciences Humaines, Dhar El Mahraz – Fès, pp 99–120
- Guerssel M (1977) Constraints on phonological rules. *Linguist Anal* 3:267–305
- Hamzah MH, Fletcher J, Hajek J (2016) Closure duration as an acoustic correlate of the word-initial singleton/geminate consonant contrast in Kelantan Malay. *J Phon* 58:135–151

- Hayes B (1986a) Assimilation as spreading in Toba Batak. *Linguist Inq* 17:467–499
- Hayes B (1986b) Inalterability in CV phonology. *Language* 62:321–351
- Hermes A, Tilsen S, Ridouane R (2021) Cross-linguistic timing contrast in geminates: a rate-independent perspective. In: *Proceedings of the 12th international seminar on speech production*, pp 52–55
- Hirata Y, Whiton J (2005) Relational acoustic invariance in the single/geminate stop distinction in Japanese. *J Acoust Soc Am* 117(4):2569–2569
- Idemaru K, Guion SG (2008) Acoustic covariants of length contrast in Japanese stops. *J Int Phon Assoc* 38(2):167–186
- Jakobson R, Fant G, Halle M (1952) *Preliminaries to speech analysis: the distinctive features and their correlates*. MIT Press, Cambridge, MA
- Jebbour A (1996) *Contraintes prosodiques et morphologie en berbère (tachelhit de Tiznit -Maroc). Analyse linguistique et traitement automatique*. Université Mohammed V, Rabat
- Jebbour A (1999) Syllable weight and syllable nuclei in Tachelhit Berber of Tiznit. *Cah Grammaire* 24:95–116
- Jessen M (1998) *Phonetics and phonology of tense and lax obstruents in German*. John Benjamins Publishing Company, Amsterdam/Philadelphia
- Kawahara S (2015) The phonetics of sokuon, or geminate obstruents. In: Kubozono H (ed) *Handbook of Japanese phonetics and phonology [Handbooks of Japanese languages and linguistics 2]*. Mouton, Berlin, pp 43–78
- Kenstowicz M (1970) On the notation of vowel length in Lithuanian. *Pap Linguist* 3:73–113
- Kenstowicz M (1994) *Phonology in generative grammar*. Blackwell, Cambridge, MA
- Kenstowicz M, Pyle C (1973) On the phonological integrity of geminate clusters. In: Kenstowicz M, Kisseberth C (eds) *Issues in phonological theory*. Mouton, The Hague, pp 27–43
- Keyser SJ, Stevens KN (2006) Enhancement and overlap in the speech chain. *Language* 82(1):33–63
- Kluender RK, Diehl RL, Wright R (1988) Vowel-length differences between voiced and voiceless consonants: an auditory explanation. *J Phon* 16:153–169
- Kohler K (1984) The phonetic explanation in phonology: the feature fortis/lenis. *Phonetica* 41:150–174
- Kubozono H (ed) (2017) *The phonetics and phonology of geminate consonants*, vol 2. Oxford University Press, Oxford
- Ladd DR, Scobbie JM, Local J, Ogden R, Temple R (2003) External sandhi as gestural overlap? Counterevidence from Sardinian. *Phonetic Interpret Pap Lab Phonol* 6:164–182
- Lahiri A, Hankamer J (1988) The timing of geminate consonants. *J Phon* 16:327–338
- Lahrouchi M (2008) A templatic approach to gemination in the imperfective stem of Tashlhiyt Berber. *Stud Afr Linguist* 37(1):21–60
- Lahrouchi M (2010) On the internal structure of Tashlhiyt Berber triconsonantal roots. *Linguist Inq* 41(2):255–285
- Leben W (1980) A metrical analysis of length. *Linguist Inq* 11:497–510
- Lehiste I, Morton K, Tatham MA (1973) An instrumental study of consonant gemination. *J Phon* 1(2):131–148
- Local J, Simpson A (1999) Phonetic implementation of geminates in Malayalam nouns. In: *Proceedings of the 14th international congress of phonetic sciences*, pp 599–612
- Louali N, Maddieson I (1999) Phonological contrast and phonetic realization: the case of Berber stops. In: Ohala JJ, Hasegawa Y, Ohala M, Granville D, Bailey AC (eds) *Proceedings of the 14th international congress of phonetic sciences*, pp 603–606
- Louali N, Puech G (1994) Les consonnes “fortes” du berbère: indices perceptuels et corrélats phonétiques. *Actes des 20^e Journées d’Etudes sur la Parole*, pp 459–464
- Malécot A (1968) The force of articulation in American stops and fricatives as a function of position. *Phonetica* 18:95–102
- Malécot A (1970) The lenis-fortis opposition: its physiological parameters. *J Acoust Soc Am* 47:1588–1592

- McCarthy J (1986) OCP effects: gemination and antigemination. *Linguist Inq* 17:207–263
- Mikuteit S, Reetz H (2007) Caught in the ACT: the timing of aspiration and voicing in east Bengali. *Lang Speech* 50(2):247–277
- Miller A (1987) Phonetic characteristics of Levantine Arabic geminates with differing morpheme and syllable structures. *Ohio State Pap Linguist Lab* 36:120–140
- Mitchell TF (1957) Long consonants in phonology and phonetics. In: *Studies in linguistic analysis*. Oxford University Press, Oxford, pp 182–205
- Muller JS (2001) The phonology and phonetics of word-initial geminates. PhD thesis, Ohio State University
- Ohala M (2007) Experimental methods in the study of Hindi geminate consonants. In: Solé MJ, Beddor P, Ohala M (eds) *Experimental approaches to phonology*. Oxford University Press, Oxford, pp 351–368
- Ouakrim O (1994) Un paramètre acoustique distinguant la gémination de la tension consonantique. *Etudes Documents Berbères* 11:197–203
- Ouakrim O (1999) A perceptual study of tenseness: some acoustic cues identifying tense vs. non-tense contrast in Berber. In: *Proceedings of the 15th international congress of phonetic sciences*. Barcelona, pp 795–798
- Payne E (2006) Non-durational indices in Italian geminate consonants. *J Int Phon Assoc* 36(1): 83–95
- Pickett ER, Blumstein SE, Burton MW (1999) Effects of speaking rate on the singleton/geminate consonant contrast in Italian. *Phonetica* 56:135–157
- Pierrehumbert J (2002) Word-specific phonetics. In: Gussenhoven C, Warner N (eds) *Papers in laboratory phonology 7*. Mouton, Berlin, pp 101–140
- Pierrehumbert JB, Clopper CG (2010) What is LabPhon? And where is it going? In: Fougerson C, Kuhnert B, D’Imperio M, Vallée N (eds) *Papers in laboratory phonology 10*. Mouton, Berlin, pp 113–132
- Pind J (1995) Speaking rate, voice-onset time, and quantity: the search for higher-order invariants for two Icelandic speech cues. *Percept Psychophys* 57(3):291–304
- Pyle C (1971) On the treatment of length in phonology. Unpublished PhD dissertation, University of Illinois
- Ridouane R (2003) Suites de consonnes en berbère: Phonétique et phonologie. Thèse de Doctorat, Université Paris 3-Sorbonne Nouvelle
- Ridouane R (2007) Gemination in Tashlhiyt Berber: an acoustic and articulatory study. *J Int Phon Assoc* 37:119–142
- Ridouane R (2010) Geminates at the junction of phonetics and phonology. In: Fougerson C, Kuhnert B, D’Imperio M, Vallée N (eds) *Papers in laboratory phonology 10*. Mouton de Gruyter, Berlin, pp 61–90
- Ridouane R, Hallé PA (2017) Word-initial geminates. In: Kubozono H (ed) *The phonetics and phonology of geminate consonant*. Oxford University Press, Oxford, pp 66–84
- Ridouane R, Fuchs S, Hoole P (2006) Laryngeal adjustments in the production of voiceless obstruent clusters in Berber. In: Harrington J, Tabain M (eds) *Speech production: models, phonetic processes, and techniques*. Psychology Press, Macquarie University, Sydney, pp 249–267
- Saib J (1977) The treatment of geminates: evidence from Berber. *Stud Afr Linguist* 8:299–316
- Stevens KN, Keyser SJ (2010) Quantal theory, enhancement, and overlap. *J Phon* 38(1):10–19
- Swadesh M (1937) The phonemic interpretation of long consonants. *Language* 13:1–10
- Taylor M (1985) Some patterns of geminate consonants. *Univ Chic Work Pap Linguist* 1:120–129
- Thurgood G (1993) Geminates: a cross-linguistic examination. In: Nevis JA, McMenamin G, Thurgood G (eds) *Papers in honor of Frederick H. Bengelman on the occasion of the twenty-fifth anniversary of the Department of Linguistics*. California State University, Fresno [Department of Linguistics], Fresno, pp 129–139
- Zeroual C (2006) Etude par transillumination des occlusives simples et géménées de l’arabe marocain. *Actes des 26^e Journées d’Etudes sur la Parole*, pp 465–469

Section II

Cultural Linguistics, Sociolinguistics, Psycholinguistics, and Pedagogy

Le corse et l'amazighe, de la minoration à l'officialisation et de la famille à l'école

6

Bouchra El Barkani

Contents

Introduction	124
Pourquoi les langues minorées/minorisées et minoritaires sont-elles menacées de mort ?	125
L'école et la sauvegarde des langues menacées	126
Gros plan sur la langue corse et son introduction à l'école	127
Le corse, d'un dialecte de l'italien à une langue autonome	127
L'introduction du corse à l'école	130
La langue polynomique: un concept pour gérer la variation linguistique	133
La théorie de l'élaboration linguistique ou la production didactique en corse	135
La langue corse: le choix d'une norme graphique	136
Variation et norme graphique: une graphie pour le corse	137
Réflexions sur la didactique de la langue corse et de sa norme graphique	138
L'amazighe au Maroc, une langue en transformation	140
L'amazighe: d'une langue minorée à une langue officielle	140
La création de l'IRCAM et l'institutionnalisation de l'amazighe	140
La langue amazighe: de l'aménagement de son corpus à son enseignement	142
Quelle norme graphique pour la langue amazighe?	142
La production didactique et l'accès de l'amazighe à l'école	143
Quelle langue amazighe enseigner à l'école primaire ?	145
Comment l'amazighe est enseigné en classe?	145
Conclusion	147
Annexes	148
Références bibliographiques	149

Abstract

En corse, la langue corse, suite à un processus d'hégémonie, a été reconnue comme un dialecte de l'italien et c'est grâce à un processus d' « individuation sociolinguistique (« *L'individuation sociolinguistique est le processus par lequel*

B. El Barkani (✉)
Institut Royal de la Culture Amazighe, Rabat, Morocco
e-mail: elbarkani@ircam.ma

une communauté ou un groupe social tend à systématiser ses différences, à les sacraliser, à les considérer comme déterminantes, à en faire un élément de reconnaissance. Elles deviennent alors des indicateurs d'identité, Z. Muljačić parle de langues par distanciation, après Kloss. » (Marcellesi 1986: 169).) » qu'il a été déclaré 'langue autonome' et langue régionale de la France en janvier 1974 date de l'extension de la loi Deixonne de 1951. Au Maroc, par contre, et après de longues années de minoration (« *La minoration (ou minorisation) linguistique est entendue ici comme la réduction, selon des degrés divers, de l'exercice sociétal normal (et donc des domaines communicationnels) d'une langue. Une réduction pouvant remonter aux origines mêmes de son émergence mais le plus souvent fruit d'une domination subie par une communauté linguistique au cours d'une plus ou moins longue période (plusieurs siècles éventuellement) et selon des modalités plus ou moins violentes.* » (Boyer 2006: 261).), voire de minorisation, la langue amazighe, a pu jouir d'un statut de « langue nationale » grâce à son introduction à l'école en septembre 2003 et en juillet 2011, elle est devenue « langue officielle » aux côtés de la langue arabe.

Entre la situation du corse et celle de l'amazighe intervient la question de la sauvegarde des langues minorées, minoritaires et régionales. Même s'il s'agit de deux langues qui ne relèvent pas du même territoire, le présent article met en parallèle ces deux situations parce qu'il s'agit de cas de langues minorées promues en langues enseignées à l'école d'où l'intérêt pour nous de voir comment ces deux langues ont évolué en terme de statut (langue minorée/langue (co)-officielle) et le rôle de l'école dans la mise en œuvre du nouveau statut des langues corse et amazighe. Cependant, en quoi consiste le rôle de l'école dans la sauvegarde des langues minorées et dans la mise en œuvre effective du changement de statut des langues corse et amazighe ? Telles sont les questions que nous allons traiter dans le présent article.

Introduction

Dans un contexte où deux langues coexistent « *lorsque l'une remplit des fonctions que l'autre ne remplit pas* » (Coyos 2004: 18), l'un des objets d'une politique linguistique est de déterminer, quelles fonctions, quelle place, souhaiterions-nous accorder aux deux langues existantes, sachant que l'avenir de chacune des langues dépendra des décisions prises dans ce cadre. Comment peut-on déterminer les fonctions d'une langue dans une situation diglossique où coexistent par exemple, une langue dominante/haute et une langue dominée/basse ?

En s'intéressant aux langues corse et amazighe et à leur introduction à l'école, nous remettons en question la situation des langues minorées voire minorisées ou les langues minoritaires dans leur combat pour survivre dans un monde où le pouvoir d'une langue donnée peut mettre fin à une autre langue. Pourquoi ces langues sont-elles menacées ? et comment l'école peut-elle les sauver ?

Pourquoi les langues minorées/minorisées et minoritaires sont-elles menacées de mort ?

Au sein de la disparité des appellations « langue minorée », « langue minoritaire », ... , la notion de « langue régionale » reste l'appellation la plus utilisée quand on parle des langues de France. (Les termes « minoration » et « minorisation » sont employés dans cet article, dans le sens de: aspects quantitatifs faisant d'une langue qui a subi un processus de minoration (moins de locuteurs qui parlent la langue) ou aspects qualitatifs faisant subir à une langue donnée, un processus de minorisation (langue sans statut défini). Ils ont été analysés par Philippe Blanchet en passant en revue les définitions données à ces termes en sociologie, en anthropologie, en sciences politiques et en sociolinguistique (Cf. Blanchet 2005).) Partant de l'exemple du français au Québec, du catalan en Espagne, *les exemples d'une défense réussie de langues en danger démontrent que la présence d'un puissant soutien institutionnel (mass média, services officiels dans la langue minoritaire, éducation bilingue) est indispensable, mais (...) elle n'est qu'un élément dans la lutte.* (Coyos 2004: 18)

En se référant à René Appel et Pieter Muysken (Cf. Coyos 2004), Anthony Lodge cite trois éléments dont la présence simultanée peut entraîner la lutte contre le transfert des langues:

- Le facteur de « statut » où *plus le statut des locuteurs et le statut de la langue minoritaire sont élevés, plus elle [la langue minoritaire] aura des chances de survivre;* (Coyos 2004: 18)
- Le facteur de « démographie » où il est question de la concentration des locuteurs de la langue ainsi, *plus les locuteurs sont concentrés géographiquement, surtout en milieu urbain, plus la langue aura des chances de survivre* (Coyos 2004: 18). Ce facteur est lié à celui de « statut » puisque la transmission de la langue, qu'elle soit corse ou amazighe, de génération en génération permet la survie de cette langue mais le fait de rester « une langue orale » les empêche de garder des traces de cette langue à l'écrit à destination des générations à venir. Ce passage à l'écrit est devenu nécessaire avec l'accès de la langue minorée, voire minorisée à des domaines dont elle a été privée d'usage comme l'école et à remplir une nouvelle fonction de « langue nationale ».
- Le facteur de la « similitude culturelle » où *plus la culture du groupe en question se rapproche de celle de la langue dominante, mieux la langue minoritaire aura des chances de survivre* (id.). Autrement dit, *si un groupe est déterminé à préserver une forte identité culturelle, il a de fortes chances de préserver sa langue* (Coyos 2004: 18).

C'est en insistant sur la « transmission de la langue menacée » dans la lutte contre le processus de « transfert des langues » d'une génération à une autre et en soulignant l'importance du regard des jeunes générations sur leurs langues qu'on peut se demander quel est le rôle de l'école dans la sauvegarde et la promotion des langues menacées.

L'école et la sauvegarde des langues menacées

Selon Christian Lagarde (2007: 16), l'école *s'est révélée un puissant obstacle à la perpétuation de la diversité linguistique* puisqu'en tant que rouage du pouvoir politique *investi du pouvoir d'inculcation d'une norme sociopolitique, socio-culturelle et sociolinguistique, [elle] produit de l'idéologie dominante*. Malgré les rôles qu'on assigne à l'école pour le sauvetage des langues menacées, *l'école n'est que l'école, c'est-à-dire un univers dont le contrôle est stratégiquement nécessaire, puisque c'est en son sein que sont délivrées de manière formalisée les compétences langagières et linguistiques, mais qui n'est guère suffisant, puisque la vie sociale n'est pas nécessairement à l'image de l'univers scolaire*. L'école peut bien, en effet, constituer un îlot de normalité linguistique que son environnement pourrait tout aussi démentir. Encore faut-il que la valeur instrumentale et symbolique de la (ou des) langue(s) dont elle pourrait constituer le sanctuaire, soit corroborée sitôt le seuil de l'école franchi. C'est là, on le sait, l'écueil majeur des RLS (Reversing Language Shift.) parmi les plus aboutis, comme ceux du Québec ou de la Catalogne (Ibid.).

Cette nécessité de l'école – qui reste insuffisante – pour sauver les langues minoritaires a été notée aussi par Raymond Renard (Cf. Lagarde 2007) qui a souligné que l'institution scolaire permet d'un côté, *d'assurer au sein des populations concernées le prestige de leur langue, évitant ainsi l'aliénation culturelle qu'implique une école étrangère au milieu*; de l'autre côté, *de favoriser la diffusion d'une norme officielle susceptible d'éviter la créolisation des langues en contact*; et d'assurer l'équilibre harmonieux des jeunes enfants aux plans psycho-cognitif, émotionnel, moral et socioculturel (Renard 2007: 30). D'après lui, (...) *l'école ne sera qu'un moyen insuffisant s'il n'est pas intégré dans une authentique politique globale d'aménagement linguistique*. Ce dernier, nécessairement plurilingue, doit prendre en compte tous les facteurs impliqués: économiques, culturels, sociaux, linguistiques, éthiques, administratifs, juridiques, éducatifs, ... (Ibid.: 32).

Devant le rôle limité et insuffisant de l'école dans le sauvetage des langues menacées, on peut souligner l'intérêt à porter aux locuteurs de la langue en question et à la politique linguistique de l'Etat dont elle relève. Donc, *Une instance particulière devrait être mise sur la conscientisation des populations*. Tous les auteurs soulignent le fait qu'une langue a peu de chance de survivre si ses locuteurs ne croient plus sa capacité à répondre à leurs besoins et sont prêts à l'abandonner pour une autre qu'ils jugent plus prestigieuse, plus apte à assurer leur promotion sociale et / ou économique (Ibid.).

Par exemple, au Maroc, nous estimons que si les locuteurs de la langue amazighe ne croient pas à la capacité de l'amazighe appris à l'école, à répondre à leurs besoins quotidiens en utilisant l'amazighe dans les administrations publiques et en dehors de l'institution scolaire, ils peuvent l'abandonner en faveur d'une autre langue qui peut être l'arabe standard ou le français à l'écrit et l'arabe marocain à l'oral.

En plus de cette conscientisation des populations, il faut aussi, selon Raymond Renard, que *l'Etat ait formulé clairement sa politique linguistique*. En particulier, le statut de chacune des langues concernées, ainsi que leurs fonctions respectives

devraient être précisés, et approuvés par la population (Ibid.). C'est pourquoi l'officialisation d'une langue jusque-là minorée lui peut assurer une large diffusion si ce caractère officiel se joint à un usage réel de cette langue.

Cependant, l'objectif primordial de l'Etat, lors de l'élaboration des politiques linguistiques doit être pensé en termes de conscientisation qui doit être poussée *jusqu'à la persuasion de la nécessité d'un plurilinguisme fonctionnel alliant identité et altérité, enracinement et ouverture, et favorisant à la fois une meilleure intégration nationale, régionale et internationale* (Ibid.: 33).

Pour des exemples concrets sur le rôle de l'école dans la sauvegarde des langues minorées/minorisées ou minoritaires, nous prenons dans cet article les exemples corse et amazighe. Comment ces langues ont-elles pu intégrer l'école ?

Gros plan sur la langue corse et son introduction à l'école

La langue corse avant d'intégrer l'école en Corse a relevé le défi de se déclarer « langue autonome » alors qu'elle était considérée comme un dialecte de l'italien. Un bref aperçu sur ce passage et sur certains concepts sociolinguistiques est exigé avant d'aborder son enseignement.

Le corse, d'un dialecte de l'italien à une langue autonome

En 2003, la Corse est devenue une *Collectivité Territoriale spécifique*. L'île est formée, de deux départements français: la Haute-Corse (no 2A) dont le Chef-lieu est Bastia et la Corse-du-Sud (no 2B) dont le Chef-lieu est Ajaccio (Cf. Annexes, Carte 1).

D'après les études consultées à ce propos (entres autres, Comiti 2005; Di Meglio 1997), le corse parlé existe, linguistiquement parlant, depuis plus d'un millénaire et jusqu'à l'annexion à la France, la Corse avait comme langue officielle « l'italien ». Et c'est *au moment du développement de l'école de la IIIe République, en français, que l'autonomie linguistique de l'île, vis-à-vis de la péninsule italienne, commence à être revendiquée* (Marcellesi 2001, p. 14). C'est en janvier 1974 que le corse a été ajouté à la liste des langues régionales de France dans une extension de la loi Deixonne de 1951.

Dans sa forme linguistique, la langue corse est formée d'un ensemble de variétés linguistiques qui se répartissent sur l'île de manière différenciée. Les linguistes ont l'habitude d'évoquer deux grandes variétés, *u cismuntincu* et *u pumuntincu*. Autrement dit, le corse du nord et le corse du sud. Par contre, Jean-Marie Comiti (2005), lui, parle de trois aires linguistiques dont le corse septentrional, le corse central, et le corse méridional. Pour ce dernier, chacune de ces aires linguistiques correspond à un « régiolecte » qui *se décompose ensuite en sous-variétés correspondant à des dialectes; ces derniers se subdivisent en entités encore plus petites, les parlers* (Comiti 2005: 92). Dans cette appartenance au groupe linguistique italo-roman, la langue corse, selon Jean Sibille (2000), est proche de

l'italien. Ainsi, les parlers du nord et de l'est de l'île sont proches des parlers ruraux de la Toscane, tandis que ceux du sud présentent de nombreux points communs avec les dialectes de l'Italie méridionale et, sur le plan phonétique, avec le Sarde.

Les dialectes de l'est de l'île de Beauté sont proches des dialectes ruraux de la Toscane et ceux du sud de l'île, des dialectes de l'Italie méridionale, notamment du Calabrais. (...) Jusqu'au XIXe siècle, le rapport entre corse et italien (toscan) était perçu comme deux niveaux d'une même langue, la lingua italiana, le corse étant la langue parlée et l'italien la langue de la culture écrite. (Fusina 1984: 84)

Si, à cause de cette parenté du corse avec l'italien, le système d'écriture corse a été adapté de l'italien officiel (*Manuel pratique d'orthographe corse* de Dominique-Antoine Geronimi et Pascal Marchetti), Jean-Baptiste Marcellesi (2001) voit dans les « indicateurs linguistiques de corsité » des traits permettant l'identification et la discrimination du corse par rapport à l'italien. Ces indicateurs linguistiques relèvent du processus d'« individuation sociolinguistique » qui a été suivi par les Corses dans le cadre de la recherche de l'autonomie de leur langue. Nous y reviendrons plus loin.

À la fin du XVIIIe siècle et jusqu'au Second Empire, s'installe en Corse une co-officialité de fait des langues italienne et française qui consacre pendant des décennies la persistance de l'ancienne culture de référence bien après la domination génoise. Selon Jean Chiorboli (2002), *cette co-officialité porte sur diverses pièces essentielles de l'administration : à commencer par le code corse dont les 14 volumes, recueil de toutes les pièces (écrits, déclarations, lettres patentes, arrêtés et règlements) publiés « dans l'île depuis la soumission à l'obéissance du Roi », qui est édité en deux langues de 1778 (premier volume) à 1790, soit plus de vingt années après la conquête de 1769. Après la révolution, il en va de même pour le Bulletin des Lois de la République, le Journal du Département de la Corse, bulletin officiel de l'époque, qui est bilingue de 1817 à 1824 [français/italien]. (Fusina 1984: 98)* Après l'annexion à la France (L'annexion de la Corse à la France a fait que l'Assemblée de Corse siège à Ajaccio, sans autre prérogative. Disponible sur: http://www.tlfq.ulaval.ca/axl/monde/Plan_du_site.htm), on assiste, en Corse, à un bilinguisme français-italien. Mais, à partir du second Empire, l'italien cède de plus en plus la place au français comme langue écrite. Ainsi, le terrain linguistique s'est trouvé conquis par le français et une littérature d'expression corse commence à s'épanouir pour déclarer, par la suite, l'autonomie linguistique du corse par rapport au français. Donc, à la place de la parenté corse/italien (diglossie stable dont le corse serait la variété basse (B) et l'italien en est la variété haute (H)) s'installe un conflit, une diglossie instable (comme l'exemple catalan/castillan) corse/français déclarant le détachement de l'italien et l'autonomie de la langue corse, reconnue depuis 1974 comme une langue régionale de France.

En se référant à la linguistique structurale, *revendiquer une parenté [du corse] avec l'italien comme le font les autonomistes de A Muvra est en partie vrai, même si la référence exclusive au toscan littéraire n'est pas recevable. (...) Claude Hagège (1987) considère le corse comme un dialecte, bien qu'il souligne prudemment que ce concept soit relativement flou (Ottavi 2007: 140).*

En parallèle, en sociolinguistique, un dialecte, selon Heinz Kloss (...) peut accéder à un statut supérieur d'une part, par la pratique systématique de l'écrit et l'extension des domaines d'usage de celui-ci (c'est le phénomène de l'élaboration linguistique). D'autre part, par la revendication de la communauté parlant ce dialecte, de la spécificité de son idiome, même en présence d'indéniables analogies avec d'autres dialectes ou une autre langue. e Il s'agit pour Heinz Kloss de « volonté populaire » (Cf. Ottavi 2007).

Donc, entre la linguistique structurale et la sociolinguistique et entre l'autonomie et l'hégémonie, la parenté corse/italien s'est-elle vue remplacée par le bilinguisme français/italien pour déclarer, par conséquent, le corse comme langue autonome (Cf. la loi Deixonne n°51-46 du 11 janvier 1951 et les décrets n°70-650 du 10 juillet 1970 et n°74-33 du 16 janvier 1974 relatif à l'enseignement des langues et dialectes locaux).

L'histoire du corse, présentée brièvement ci-dessus, relate les revendications des différentes associations corses qui ont abouti, après l'intervention du pouvoir politique, à une reconnaissance du corse en tant que langue et en tant que partie prenante des langues enseignées à l'école. Alors, du point de vue de la linguistique structurale, le corse est un dialecte de l'italien tandis que la sociolinguistique consacre l'autonomisation de cette langue.

L'accession du corse au statut de langue surtout de langue enseignée est, selon Alain Di Meglio (2002, juin), le résultat d'un cheminement historique particulier où *l'émergence de sa conscience n'est pas le résultat d'une déclaration politique à partir d'un centre mais d'une différenciation dont nous décrivons le processus à travers diverses acceptions de diglossie qui détermine un rapport hiérarchique entre deux variétés linguistiques ou deux langues dans une même communauté. Le concept de Ferguson (diglossie) revisité sensiblement par Fishman vient expliquer la stabilité et la neutralité du rapport langue/dialecte jusqu'aux années 70. La conception catalaniste introduira le paramètre de l'évolution vers la disparition de l'une des deux langues qui entrent alors en conflit. (...) Face à cette pression, l'école s'adapte par la permissivité du politique et ouvre progressivement un espace où peut se développer un nouveau savoir: la langue corse* (Ibid.).

Dans le flux de la discussion qu'a suscitée la situation linguistique en Corse, à prendre la langue comme une langue standard ou un ensemble de variétés à unité abstraite, un nouveau concept a été annoncé par Jean-Baptiste Marcellesi en 1983, celui de « langue polynomique ». Cette polynomie baptise une plurinorme dans laquelle *se développent des attitudes tolérantes vis-à-vis de la variation dialectale. Cette polynomie renvoie bien à des « attitudes » et non, comme beaucoup ont tendance à le penser, à la seule variation linguistique. Sinon toutes les langues seraient polynomiques dans la mesure où elles présentent toutes un degré important de variation* (Comiti 2005: 96). C'est ce concept de « polynomie » qui a donné un cadre à l'enseignement du corse.

L'introduction du corse à l'école

D'après Jean Sibille (2000), le mouvement culturel corse, dont fait partie *A Muvra*, journal corse, organe du *Partitu Corsu d'Azione*, animé par Petru Rocca qui réclamait des thèmes de revendication à forte coloration politique (Cf. Lafont and Ernest 1982), privilégie l'apprentissage, par l'élève, de la variété parlée dans le lieu où il habite tout en lui apprenant les autres formes dialectales pour en développer une connaissance passive.

Le mouvement culturel corse n'a pas cherché à imposer une langue unifiée à l'ensemble de l'île, préférant accepter « la langue corse telle qu'elle est, dans cette dialectique de l'un et du multiple qui est celle de la vie » (F. Ettori). Les linguistes corses parlent de « langue polynomique » : son enseignement est fondée sur chaque variété lexicale; l'enfant apprend d'abord le parler du lieu où il habite et s'initie peu à peu à d'autres variétés; pour finir par acquérir une connaissance passive de l'ensemble des parlers de l'île (Sibille 2000: 32–33).

Cette inclination vers le développement d'une « langue polynomique » basée sur les variétés dialectales préexistantes vise en parallèle, à permettre à l'apprenant d'acquérir une connaissance passive sur les dialectes autres que le sien. Cette connaissance peut être mise en œuvre pour comprendre l'autre mais elle n'oblige pas l'apprenant à parler un dialecte autre que le sien. L'objectif donc est de développer l'intercompréhension.

La restriction de l'enseignement/apprentissage de la langue corse lie l'émancipation de cette langue dans la société à la motivation des élèves et des familles qui, seuls, peuvent choisir de bénéficier ou non de ces cours. Autrement dit, il faut qu'ils trouvent dans le suivi des cours de langue corse une motivation pour en bénéficier. Mais, quand l'élève découvre qu'il ne retrouvera plus cette discipline (discontinuité) dans son cursus, il se dispense souvent de ces cours.

Pour favoriser la motivation des élèves pour apprendre le corse, il faut le faire *par une politique résolument tournée vers un bilinguisme scolaire généralisé et officiel (...) [qui] donnera à la langue corse les moyens d'investir tout l'espace social, sinon elle risque fort d'endosser à jamais la seule fonction de discipline scolaire dans une école qui lui servira d'unique refuge* (Sibille 2000: 87). La deuxième difficulté est liée à la généralisation nécessaire de l'enseignement du corse au primaire. En l'absence de cette généralisation, les Corses soulignent la nécessité de multiplier les sites bilingues français/corse et de multiplier les sites de séjour qui s'engagent à assurer un enseignement du corse par immersion, par un bain linguistique (Cf. Di Meglio and Cortier 2005).

Selon Jean-Marie Comiti, si l'objectif de l'école est de sauvegarder et valoriser la langue corse, le traitement de cette langue peut être abordé sur deux plans: la politique linguistique et son statut de discipline scolaire. Si pour lui, la politique linguistique vis-à-vis du corse à l'école est discriminatoire, son statut de « discipline scolaire » n'offre pas à la langue corse un usage en dehors de l'école. Donc, l'institution scolaire reste l'enceinte où le corse (en tant que langue) naît et se développe (la métaphore du syndrome du bocal de Jean-Marie Comiti (2005)).

L'analyse sociolinguistique corse s'est appuyée, d'une part, sur le concept de la diglossie tel qu'il a été mis en place par Charles Ferguson et développé par Joshua Fishman (Cf. Comiti 2005) et sur la notion de glottophagie lors de la francisation (du début du siècle jusqu'aux années soixante-dix). D'autre part, elle s'est basée sur les concepts de « l'individuation sociolinguistique » et de « l'élaboration linguistique » de Heinz Kloss (Cf. Di Meglio 1997). Il s'agit de concepts fondés, d'un côté, sur la notion d'*Ausbau* ou « le développement de la langue » et dont le fondement est la volonté populaire pour « l'élaboration linguistique », et de l'autre côté, sur l'idée d'autonomie linguistique dans des situations d'assimilation pour le concept d'« individuation sociolinguistique ». Et enfin, en se basant sur le concept de « polynomie » instauré par Jean-Baptiste Marcellesi, des réflexions sur l'enseignement de la langue corse basé sur « l'intolérance dialectale » et le respect de la variation ont été déclenchées.

d'après les sociolinguistes corses et catalans, tout conflit linguistique entre langues dominées et langues dominantes mène vers soit une assimilation ou une normalisation. Le passage du corse, comme langue dominée par le français et l'italien, de « dialecte » à « langue autonome », est le résultat, selon Alain Di Meglio et Claude Cortier (2005), non pas d'un processus officiel et/ou scolaire, mais d'un *cheminement d'auto-reconnaissance* (Ce que Jean Baptiste Marcellesi a nommé le processus « d'individuation linguistique ».) *par ses propres locuteurs dans un contexte revendicatif et politique issu du renouveau des cultures régionales en Europe à partir des années soixante-dix* (Di Meglio and Cortier 2005). S'inscrivant dans le même champ opératoire, les concepts d'« individuation sociolinguistique », de « reconnaissance-naissance » (« *La reconnaissance-naissance est un phénomène externe par rapport à l'individuation sociolinguistique qui est le fait de la communauté concernée. L'une et l'autre participent au processus qui fait d'une variété une différence.* » (Marcellesi 1986: 180)), et de « polynomie » ont assuré la base du projet de l'enseignement de la langue corse. Et si les concepts de « -reconnaissance-naissance » et d'« individuation sociolinguistique » visent l'autonomie de la langue, le concept de « langue polynomique » vise le mode d'existence de la langue et son rapport à la norme. De ce fait, en présence de plusieurs idiomes, seule la conscience linguistique est capable de nous dévoiler s'il s'agit de plusieurs langues différenciées ou d'une seule langue avec différentes variantes.

Il ne faut pas dire, dès qu'il y a variation, dès qu'il y a un certain nombre de variations, dire que ce sont des langues différentes. Le problème c'est celui d'une certaine conscience linguistique de ce qui est la même langue, ce qui n'est pas la même langue. (Marcellesi 2001: 14)

Selon Jean-Marie Comiti (2005), il apparaît clairement qu'*une distanciation d'ordre linguistique, sociolinguistique et psycholinguistique, a été opérée et que le corse est devenu aux yeux des Corses une langue autonome qui n'entretient avec les autres langues romanes que les affinités dues à l'origine latine commune* (Comiti 2005: 58). C'est sur cette conscience linguistique que Jean-Baptiste Marcellesi s'est basé, en 1984, dans son enquête menée sur le territoire corse, pour mettre à l'épreuve

un processus abordé par Heinz Kloss et repris par lui. Il s'agit de « l'individuation sociolinguistique » qui, comme nous l'avons défini au début de ce texte, a pour but, une auto-reconnaissance du corse comme langue autonome et non comme dialecte de l'italien. C'est par ce processus que, selon Alain Di Meglio et Jean-Marie Comiti (1999), *une communauté définit sa langue sur la base d'indicateurs linguistiques d'identité et la déclare autonome des systèmes voisins* (Di Meglio and Comiti 1999: 62–63). Ces indicateurs linguistiques d'identité sont des critères qui permettent à la communauté corse de déclarer sa langue « autonome ».

Contrairement à la domination dont l'exemple de la minoration du catalan sous le régime franquiste, le processus d'hégémonie, qui s'oppose au processus de la « conscience linguistique », *s'accompagne d'une certaine forme de conviction et de consentement* (Marcellesi 1986: 167) comme, par exemple, la subordination des dialectes italo-romans à l'italien. Cette subordination est la « satellisation » ou le résultat sociolinguistique des rapports d'hégémonie. Il s'agit de *systèmes linguistiques généralement proches mais parfois nettement différents du système dominant [qui] s'organisent comme s'ils dépendaient de lui, devaient s'effacer devant lui, comme s'ils étaient régis par lui* (Ibid.: 169). Ainsi, le couple antithétique individuation/satellisation s'oppose à la reconnaissance-naissance et oppose ainsi, le sociolinguiste au linguiste. Autrement dit, si le linguiste *essaie de classer les langues à partir de certains pôles considérés comme imposés par l'évidence et mesure ensuite les distances structurelles, avec l'intention de ne considérer comme « langues autres » que celles qui se distinguent de manière systématique, sur des points importants* (Ibid.: 170), le sociolinguiste, par contre, *prend en compte une situation globale, et notamment l'attitude que la communauté concernée observe vis-à-vis des autres communautés sociolinguistiques, la valeur des modèles qu'elle leur reconnaît ou ne leur reconnaît pas, l'acceptation ou non de leurs instances normalisatrices* (Ibid.: 170). Ainsi, *Dès lors que la communauté manifeste avec persévérance la volonté d'être autre, de nommer sa langue d'un nom particulier qui ne soit pas un sous-ensemble d'une autre langue, qu'elle établit sa propre stratégie et au besoin développe ses propres mythes fondateurs, on entre dans un processus d'individuation sociolinguistique qui va prendre appui sur les différences plus minimes avec autrui, il ne sert à rien à ce moment-là d'opposer à ce sentiment les proximités structurelles ou l'intercompréhension* (Ibid.: 170).

Parallèlement au concept d'« individuation sociolinguistique », la reconnaissance-naissance – qui reconnaît à un idiome son autonomie par rapport à d'autres idiomes – est un processus dont *nous entendons les décisions glottopolitiques qui s'appliquent à un système linguistique qu'on classait antérieurement comme dépendant d'un autre, les différences linguistiques étant minimisées, et qui érigent ces différences en écarts significatifs, procédant ainsi à la connaissance de ce qui existait déjà sur le terrain et donnant naissance par déclaration solennelle à une nouvelle langue* (Ibid.: 280). Donc, en complémentarité avec « l'individuation sociolinguistique » qui émane de la communauté qui se base sur la conscience linguistique pour déclarer le corse différent de l'italien et ainsi « langue autonome », la reconnaissance-naissance émane d'une décision solennelle déclarant un idiome présent sur le terrain comme une langue autonome. Autrement dit, à une intervention « interne » s'oppose une autre intervention de type « externe ».

Par conséquent, en partant du slogan « Diversità faci ricchezza » ou « la diversité constitue une richesse » de la revue corse *A Spannata* qui défend les parlers du sud, on symbolisait, par ce slogan, selon Alain Di Meglio, en plus de la position en faveur de la prise en compte de la variation géographique, une réaction fondée sur l'affirmation de l'existence d'une hégémonie linguistique et culturelle du nord de la Corse.

On arrive ainsi à l'expression de Fernand Etori (1975) et Jacques Fusina (1984, 2003) sur « la dialectique de l'un et du multiple » justifiant le choix d'une norme plurielle qui s'est imposée au corse avec le manuel pratique d'orthographe « *Intricciate è cambiarine* » (1971) de Geronimi et Marchetti (Cf. Fusina 1984; Di Meglio 1997), d'où l'émergence du concept de « langue polynomique » comme initiative de gestion de la variation.

La langue polynomique: un concept pour gérer la variation linguistique

La variation, un obstacle à l'intercompréhension ou une richesse à cultiver ? Telle est la question à poser lorsqu'il s'agit d'enseigner/apprendre une langue qui était, pour longtemps, « minorée » et qui devient une discipline scolaire dont l'histoire fait remonter son statut de « langue » à son existence sous forme de plusieurs variantes dialectales. Comme l'a noté Alain Di Meglio, *la variation interne d'une langue minorée constitue de façon récurrente un obstacle à son enseignement* (Di Meglio 1997: 151) et dans le cas du corse *l'unification et la normalisation orthographique ont très tôt été posées comme condition sine qua non de son accès à l'enseignement dans l'argumentaire glottopolitique* (Ibid.: 151). La normalisation, telle qu'elle a été définie par les Catalans, est une « normativisation » (élaboration d'une codification de plus en plus raffinée) en extension sociale (diffusion dans tous les domaines d'emploi). Elle est la base de l'enseignement du corse.

Nous entendons par « normalisation » l'ensemble formel de prescriptions auquel se réfèrent les locuteurs pour un usage correct de la langue; par « standardisation » l'élaboration d'un usage commun, un standard, susceptible d'assurer la communication pour le plus grand nombre; par « codification » le système de règles qui concerne tout particulièrement l'orthographe. (Comiti 2005: 90)

En Corse, il n'existe pas, selon Jean-Marie Comiti (2005), une institution officielle qui se charge de la normalisation de la langue corse. Mais, *c'est une tâche que les politiques délèguent implicitement à l'institution scolaire. C'est en effet le lieu idéal où s'élaborent et se transmettent, selon les représentations les plus largement partagées, les règles du bon usage de la langue* (Comiti 2005: 89).

La décision d'enseigner le corse a été, d'après Jean-Baptiste Marcellesi (2001), une décision politique. Et entre cette décision politique d'enseigner la langue corse et le choix de la norme à enseigner, il fallait penser à débloquer une situation où chacun voulait imposer sa variété, en prenant cette langue comme un ensemble de variétés

dialectales constituant sa richesse. Comme disait Fernand Etti (1975) « la richesse est dans la diversité ». Ainsi, le concept de « langue polynomique » a permis d'inclure la diversité (plurinorme) sans privilégier une variété vis-à-vis d'une autre.

Il y avait à l'époque un conflit (...) sur la forme du corse qu'il fallait enseigner (...) il y avait une décision d'enseigner le corse (...) mais on pensait à la norme, et tout était bloqué puisque chacun voulait imposer sa variété. On est arrivé à la conclusion qu'il fallait débloquent la situation et on m'a demandé si ça pouvait fonctionner avec une vision scientifique sérieuse (...) à Rouen, nous avons discuté de manière que ces langues soient conçues comme un ensemble de variétés constituant un trésor commun où chacun puise selon ses origines, ses goûts, et éventuellement sa stratégie de communication. (Marcellesi 2001: 22–23)

Dans le flux de la discussion qu'a suscitée la situation linguistique en Corse, à prendre la langue comme une langue standard ou un ensemble de variétés à unité abstraite, un nouveau concept a été annoncé par Jean-Baptiste Marcellesi en 1983, celui de « langue polynomique ». Cette polynomie baptise une plurinorme dans laquelle *se développent des attitudes tolérantes vis-à-vis de la variation dialectale*. Cette polynomie renvoie bien à des « attitudes » et non, comme beaucoup ont tendance à le penser, à la seule variation linguistique. Sinon toutes les langues seraient polynomiques dans la mesure où elles présentent toutes un degré important de variation (Comiti 2005: 96).

D'après Jean-Marie Comiti (2005), le corse correspond au concept de « langue polynomique » parce que *les différentes variétés qui le composent ne font pas l'objet d'une hiérarchisation verticale, contrairement au modèle français où une sur-norme officielle règne au-dessus d'autres normes réputées inférieures. Les variétés dialectales du corse sont toutes considérées comme étant d'égale valeur linguistique et sociale* (Ibid.: 96–97). Ainsi, la « hiérarchisation horizontale » où tous les dialectes se retrouvent sur le même pied d'égalité (exemple des dialectes corses), s'oppose à une « hiérarchisation verticale » où une « sur-norme » règne en minorant d'autres idiomes comme le cas de la France où le dialecte de l'île de France (dialecte d'oïl) s'est substitué aux autres dialectes présents sur le territoire français pour bâtir une « langue française ».

La polynomie est, selon Jean-Marie Comiti (2005), *une forme de démocratie linguistique véritable qui favorise et encourage la reconnaissance et la légitimation de la différence dialectale* (Ibid., p. 97). Cette reconnaissance de la différence dialectale favorise une tolérance interdialectale qui est devenue l'un des éléments fondamentaux du concept de « langue polynomique » proposé par Jean-Baptiste Marcellesi lors du congrès international de linguistique et philologie romanes qui avait eu lieu en 1983 à Aix-en-Provence. Sa réflexion polynomique a émergé en 1980 dans un article paru dans le n° 14 d'*Etudes corses* intitulé « Pour une approche sociolinguistique du corse ». Ce concept de « langue polynomique » est apparu dans le cadre d'une prise en considération du mode d'existence des langues dont, selon Jean Baptiste Marcellesi, l'unité est abstraite et résulte d'un mouvement dialectique. Ce concept de « polynomie » va se révéler très utile dans la gestion de la variation

dialectale corse que les insulaires n'appréhendent plus comme une entrave à l'intercompréhension ni comme un obstacle au « statut de langue ».

Il s'avère que les variétés dialectales corses apparaissent comme d'égale valeur (linguistique et/ou sociale) et que rien ne justifiait que l'une d'entre elles fût choisie par quelque instance que ce soit pour jouer le rôle de norme unique. La conception d'une norme plurielle légitimée par l'ensemble de la communauté corse est à la base de la polynomie qui fait par ailleurs l'objet d'une formation originale des enseignants et d'une pédagogie adaptée au contexte d'une véritable démocratie linguistique. (Ibid.: 57)

Le fait d'accepter la langue comme une somme de variétés fondant sa richesse repose, selon Alain Di Meglio (1997) sur l'institution d'une « plurinorme » à l'écrit en maintenant, une codification graphique commune. Et c'est dans le cadre de cette « plurinorme » que la langue sera équipée (la théorie de l'élaboration didactique).

La théorie de l'élaboration linguistique ou la production didactique en corse

D'après Alain Di Meglio et Claude Cortier (2004: 185), « l'élaboration linguistique » est un processus d'équipement et d'adaptation d'une langue jusque-là figée dans le temps et confinée aux registres de la famille et de l'intimité des communautés urbaines et villageoises. Forcée aussi par Heinz Kloss en 1978, la théorie de l'élaboration linguistique vient à l'encontre des stratégies assimilationnistes. Elle s'appuie, principalement, sur le découpage de l'espace du langage en domaines d'usage définis chez Heinz Kloss dans: l'école, l'église, le parlement, la presse, le cinéma, la littérature, la radio. La langue, d'après sa présence et sa qualité dans quelques-uns de ces domaines, se développe selon un grade qui croit entre le dialecte (normal dialekt) jusqu'à la langue élaborée (ausbauprache). Žarko Muljačić (cité par Marcellesi 1986) parle, lui, d'un schéma qui va de « langues au début de leur élaboration » à « langues en cours d'élaboration » pour arriver à des « langues élaborées ». Ce schéma efface, d'après lui, le bond qualitatif qui va de « dialecte » à « langue » (Cf. Di Meglio 1997). Pour le domaine de l'école, l'ensemble des outils issus de la production didactique constitue une manifestation de la théorie de l'élaboration linguistique. Puisque l'insertion de la langue dans l'enseignement est la pierre angulaire de l'élaboration linguistique, comment s'effectue-t-elle ? La progression de l'élaboration tend à transformer la langue en instrument de communication. Donc, cette progression se situe sur une échelle dont l'objectif est une éducation bilingue (système d'enseignement bilingue français/corse).

D'après Alain Di Meglio et Claude Cortier (2004), la notion très générale d'élaboration linguistique se traduit par un parcours de développement/adaptation jalonnée par l'institution qui offrira par le jeu des circulaires et instructions officielles, plusieurs possibilités et moyens en ce sens:

- la codification s’affine en 1971 avec le *Manuel pratique d’orthographe* de Pascal Marchetti et Dominique-Antoine Geronimi et propose une norme écrite souple qui réussit à tenir compte de la variété interne du corse tout en postulant son unicité linguistique.
- la littérature se dote d’une première anthologie et propose, par le CRDP (Centre Régional de Documentation Pédagogique.) fondé en 1977, les premiers manuels didactiques pour les classes de lycées. Les instruments didactiques s’étoffent à mesure que l’on avance dans les décennies 80 et 90 jusqu’à couvrir l’ensemble du système éducatif scolaire.
- l’université de Corse, créée en 1981, affine l’appareil conceptuel autour d’une approche de la langue qui conjugue unicité et multiplicité (concept de « langue polynomique »).
- et par son utilité [le concept de « langue polynomique »] pour la légitimation des langues minorées, le concept va connaître un réel succès et servira de base aux premiers discours dans la formation initiale lors de l’avènement des IUFM en 1990 (Comiti 1999). La production didactique dans les associations a été supplantée par les structures publiques mises au service du corse dont l’assemblée territoriale de Corse relayée par le CRDP qui s’affirme comme un lieu centralisateur de la production didactique de manuels, documents et méthodes scolaires. Et partant de ce fait, il est connu comme un lieu privilégié de normalisation linguistique en l’absence d’un centre de recherche chargé de l’aménagement de la langue corse.

La minoration a été mise en question dans le cas corse en présence des revendications militantes et des politiques linguistiques qui en résultent pour veiller sur sa promotion en faisant partie des langues minorées qui *frappent aujourd’hui avec plus ou moins d’insistance à la porte de l’école demandant réparation des injustices de traitement que l’histoire leur a fait subir* (Bentolila and Germain 2005). Et c’est en se fixant une norme graphique que ces langues minorées peuvent accéder à l’école et fuir, par conséquent, ce processus de minorisation.

La langue corse: le choix d’une norme graphique

En tant que langue minorée, le corse est constitué de variétés dialectales et cette variation fait obstacle, comme nous l’avons noté plus haut, à son enseignement/apprentissage. Alors, devant la problématique: « quel corse écrire et enseigner ? », l’unification et la normalisation de la langue corse s’impose et, par conséquent, sa normalisation orthographique devient une exigence.

Comme l’a noté Alain Di Meglio (1997), *le corse, dans son histoire, n’a jamais été une langue enseignée dans un système éducatif organisé (jusqu’au siècle dernier, seule la transmission orale a assuré sa pérennité) et moins encore une langue d’enseignement. Son système d’écriture n’a donc pas subi l’œuvre de stabilisation normative que peut assurer une longue tradition scolaire et littéraire* (Di Meglio 1997: 2).

Ainsi, après l'étape de la reconnaissance-naissance de la langue corse basée sur « l'individuation sociolinguistique » (Cf. Marcellesi 2001) et dans son passage de « langue vernaculaire » à vocation orale à « discipline scolaire » (Cf. Comiti 2005), la langue corse s'est trouvée en confrontation avec l'un des problèmes relatifs à sa diffusion et ainsi, à son adaptation à la vie moderne dont l'accès à l'école. Dans cet accès de la langue corse à l'école, c'est la question de la norme graphique qui se pose.

Variation et norme graphique: une graphie pour le corse

Dans une esquisse typologique des langues, Miguel Siguán et William Mackey (1986: 113), ont présenté les langues non écrites comme les *langues qui ne sont pas totalement normalisées et qui ont été étudiées scientifiquement, dont on peut dire par conséquent que leurs normes grammaticales et leur vocabulaire sont connus, et qui peuvent être facilement écrites en adaptant les conventions phonétiques et graphiques d'une langue similaire*. Le choix d'une orthographe pour le corse par exemple, a été basé sur un processus d'adaptation graphique d'une langue similaire (en l'occurrence le toscan écrit) et qui a évolué vers des formes graphiques spécifiques fondant l'autonomie de la langue corse (Cf. Di Meglio 1997; Chiorboli 2002).

Comme toutes les langues qui accèdent à l'écrit, le système de transcription utilisé à l'origine est celui d'une langue autre choisie comme modèle. C'est ainsi que les premiers textes écrits en corse se sont appuyés sur la graphie toscane, c'est-à-dire celle de la langue écrite dominante en Corse au début du dix-neuvième siècle (Chiorboli 2002: 142).

Selon Jean Chiorboli (2002: 100), *les premières tentatives d'écriture du corse, au XIXe siècle principalement, se sont appuyées sur la base italienne. Par la suite, tantôt par volonté de simplification, tantôt par désir d'autonomisation par rapport aux graphies italiennes, tantôt par influence du français où à cause de pratiques patoisantes, les propositions diverses se sont multipliées conduisant parfois à une situation anarchique (Ibid.). En 1971, l'ouvrage Intricciate è cambiarine de Dominique-Antoine Geronimi et Pascal Marchetti proposa, selon Jean Chiorboli (2002), un système d'orthographe et d'orthoépie qui emporta l'adhésion du plus grand nombre et constitua dès lors la référence dans son domaine, à quelques détails près (Ibid.). Dans ce Manuel pratique d'orthographe corse – « Intricciate è cambiarine » -, Dominique-Antoine Geronimi et Pascal Marchetti se sont basés, dans l'adaptation graphique de l'italien au corse, sur un texte du début du XIXe siècle extrait d'une œuvre rédigée en italien. Ce texte intitulé « U Serinatu discappinu », est extrait de l'œuvre de l'écrivain Salvatore Viale intitulée La Dionomachia (la guerre de l'âne) (1817). Rédigé, d'après Alain Di Meglio (1997) dans une variété « villageoise » (région bastiaise, vu que l'écrivain est un bastiais)). Ce texte servit d'illustration dans le manuel (Intricciate è cambiarine) pour mettre en exergue les principales modifications, apportées entre les deux époques (Cf. Di*

Meglio 1997). Ces modifications apportées par Dominique-Antoine Geronimi et Pascal Marchetti au texte d'origine pour adapter la graphie italienne au corse mettent en valeur les traits graphiques qui caractérisent le corse le distinguant ainsi, des autres codes. Il s'agit des traits « d'individuation sociolinguistique » de Jean-Baptiste Marcellesi évoqués précédemment. Ces traits graphiques spécifiques à la langue corse servent à noter les sons qui sont, pour Jean Sibille, inconnus de l'italien standard pris comme référence pour l'élaboration de la norme graphique corse.

L'orthographe actuelle du corse a été élaborée par les pédagogues à partir de l'orthographe italienne, en l'adaptant avec le maximum de cohérence possible, de façon à noter des sons inconnus de l'italien standard. Elle met en valeur une certaine unité de la langue, tout en respectant la variété dialectale (Sibille 2000: 33).

Le *Manuel pratique d'orthographe corse* donne l'exemple d'une « plurinorme » où les auteurs (Dominique-Antoine Geronimi et Pascal Marchetti) ont proposé un ensemble de variantes possibles pour: d'une part, ne pas occulter les traits principaux des différentes variétés du corse et d'autre part, pour définir la langue corse comme la somme de tous les parlers qui se distinguent sur le territoire corse par de minces variations.

Dépassant les modèles des grandes langues ayant unifié leur écrit, Marchetti/Geronimi ont proposé un ensemble de variantes possibles dans le but de ne point occulter les traits principaux des différentes variétés du corse. Plutôt qu'à une norme écrite, nous avons ici affaire à un cadre normatif ou une plurinorme qui prend une première définition dans la déclaration « langue corse, la somme de tous les parlers, distingués entre eux par de minces variantes, qui sont utilisées sur le territoire de l'île de Corse » (Di Meglio 1997: 121).

Si cette « plurinorme » résout les problèmes liés au choix d'un alphabet corse, on peut se demander quelle didactique convient d'une part, pour enseigner une langue polynormique et d'autre part, pour enseigner une écriture « plurinorme ».

Réflexions sur la didactique de la langue corse et de sa norme graphique

Dans une réflexion sur la didactique de la langue corse, la spécificité de ce projet consiste dans l'objet d'enseignement qui est « une langue polynormique » basée sur « la démocratie linguistique » ou « l'intertolérance dialectale ». Cette variété dialectale est constituée de plusieurs parlers qui doivent être pris en compte dans l'élaboration d'une didactique de la langue. Au niveau de l'écrit, cette variation du 2^{ème} degré si on peut dire (la variation de la langue en dialectes en est le 1^{er} degré tandis que la variation du dialecte en parlers est le 2^{ème} degré dans cette hiérarchie), a été, selon Alain Di Meglio (1997), réduite pour *approcher la langue plutôt que les parlers*. Eviter les éléments spécifiques à une variété très localisée note la spécificité de l'enseignement de l'écrit en langue corse où on se réfère aux variétés dialectales en donnant un ensemble fixe de formes possibles à l'écrit (trois formes et rarement

au-delà) sans se limiter à en donner une forme unique. Alors, comme dans l'exemple français clé/clef, dans le corse écrit, on prévaut « l'intertolérance » dialectale pour donner les formes possibles à l'écrit en se référant à ses dialectes.

Ainsi, l'enseignement du corse développerait une tendance originale à l'ossification dans la mesure où elle donne aux éléments écrits, non pas une forme unique, mais un ensemble fixe de formes possibles qui peut aller jusqu'à trois ou rarement au-delà. Ce que fait exceptionnellement le français pour le mot clé/clef, le corse tend à le systématiser, non dans l'impossible objectif de refléter l'ensemble de la variation dialectale mais dans ce qui nous apparaît comme une représentation symbolique de celle-ci. (Ibid.: 136)

Dans le cadre de la *Commissione Academica Lingua è Cultura corsa* (commission académique de langue et culture corses), un travail a été effectué, en 1985, par l'administration scolaire, à propos des tolérances graphiques. Ce travail est intitulé « Elementi di cunversione dialettologica » (éléments pour une conversion dialectologique) (Cf. Di Meglio 1997). L'objectif de ce travail est de veiller à la construction d'une norme graphique, d'un côté, pour satisfaire un désir de construire un système normé, stable et commun à toute la Corse. Et, de l'autre côté, pour permettre la diffusion et la compréhension, par les Corses, des textes issus de ce travail.

Nous pouvons d'ores et déjà avancer l'hypothèse de la naissance de ce phénomène par la conjonction de deux impératifs dans la construction de la norme écrite: 1)-un désir de se reconnaître dans l'écrit par une somme de marqueurs locaux donnés, 2)-un désir de construction d'un système normé stable et commun à toute la Corse dans un but de grande circulation des textes, de communication nécessaire à une langue enseignée et/ou d'enseignement qui tend à réduire la fragmentation dialectale pour un système stable et fixe; « approcher la langue plutôt que les parlers » est pour OTTAVI (1992: 53) une préoccupation fondamentale que doit avoir le maître dans la classe de corse intégré. (Ibid.: 137)

Ainsi, entre l'élimination des traits graphiques de variation très localisée (au niveau des parlers) et la « conversion dialectologique », le choix d'une « plurinorme » fixant le cadre de la variation et notant la spécificité de l'enseignement du corse est une spécificité à prendre en compte lors de l'élaboration didactique en fixant les possibilités de la variation normée. Donc, *la table de conversion dialectologique donne déjà un cadre assez restreint de possibilités de variation normée* (Ibid.: 139). Cependant, l'enseignement du corse dans sa conception polynomique est différent de l'enseignement du français ou des autres langues vivantes, une nouvelle approche à laquelle les élèves ne se sont pas habitués, ce qui poussait les parents à intervenir sur les cahiers de leurs enfants pour corriger la forme de certains écrits qui ne correspondent pas à leur variété linguistique. Mais, cet enseignement est reçu, d'après Alain Di Meglio et Jean-Marie Comiti (1999), comme une ouverture sur les autres variétés.

Cette approche corse de gestion de la variation et de la norme dite graphique d'une langue minorée promue en langue enseignée nous mène à interroger la question amazighe au Maroc entre la variation et l'élaboration d'une norme graphique suite à son introduction à l'école.

L'amazighe au Maroc, une langue en transformation

Parler de l'enseignement de l'amazighe au Maroc suscite une question primordiale : *l'amazighe, langue enseignée-apprise à l'école, quoi enseigner, par qui, et comment ?* Pour y répondre, nous allons présenter d'abord, l'amazighe au Maroc, son passage de 'langue minorée' à 'langue nationale' puis 'langue officielle'. Ensuite, nous nous focaliserons sur son introduction à l'école. L'objectif est de présenter un état des lieux de l'enseignement de l'amazighe en tant que langue minorée promue au statut de langue officielle pour se focaliser sur le rôle de l'école dans la survie et la promotion d'une langue minorée.

L'amazighe: d'une langue minorée à une langue officielle

La langue amazighe est une langue orale depuis des siècles. Elle est présente sur le territoire marocain sous forme de variantes linguistiques (tarifit, tamazight, tachelhit) (Cf. Annexes, Carte 2) étendues sur le territoire. Elle a été introduite à l'école marocaine en septembre 2003. Sur le plan vertical (cycles de l'enseignement), elle est enseignée au primaire de la 1^{ère} à la 6^{ème} année et sur le plan horizontal (écoles ciblées), elle fait partie du cursus scolaire de plusieurs écoles sans que son enseignement soit généralisé à l'échelle nationale. En juillet 2011, elle a changé de statut de « langue nationale » à « langue officielle » reconnue aux côtés de la langue arabe, dans la Constitution marocaine, comme l'une des deux langues officielles du pays. Son enseignement a nécessité l'aménagement de son corpus. Le premier pas franchi dans son aménagement a été le choix de la graphie avec laquelle cette langue sera transcrite pour qu'elle soit une langue enseignée, apprise et diffusée et c'est la graphie tifinaghe qui a été choisie en février 2003 pour écrire la langue amazighe. La constitutionnalisation de l'amazighe a été précédé de son institutionnalisation avec la création de l'Institut royal de la culture amazighe (désormais IRCAM).

La création de l'IRCAM et l'institutionnalisation de l'amazighe

C'est en 2001 que sa Majesté le Roi Mohammed VI a annoncé dans son discours du 17 octobre la nécessité de promouvoir l'amazighe, composante du patrimoine national (Cf. Discours royal d'Ajdir, 17 octobre 2001).

Déclarée comme une responsabilité nationale, la promotion de la langue et culture amazighes a fait l'objet de la création, par sa Majesté, de l'institut qui prendra en charge cette mission. Il s'agit de l'Institut royal de la culture amazighe, institution à visée consultative visant la sauvegarde et la promotion de la langue et culture amazighes. Pour accomplir sa mission, l'IRCAM doit collaborer avec plusieurs institutions gouvernementales chargées des secteurs de l'éducation, la formation, les médias, ... (Cf. Article 2, Dahir 2001)

Les missions dévolues à l'IRCAM sont déclinées dans l'article 3 du Dahir le créant et l'organisant (Cf. Article 3, Dahir 2001).

Ces missions ont fait de l'IRCAM, une institution de référence en matière de recherche en amazighe et sur l'amazighe. Une place qui s'est renforcée avec l'officialisation de l'amazighe en juillet 2011 et l'adoption en août 2019 des lois organiques relatives, respectivement, à la mise en œuvre du caractère officiel de la langue amazighe et la création du Conseil national des langues et de la culture marocaine. *La langue amazighe, entre la constitution et la loi organique relative à la mise en œuvre de son caractère officiel*

Dix ans après l'institutionnalisation de l'amazighe et la création de l'IRCAM, la langue amazighe est devenue langue officielle du Maroc aux côtés de la langue arabe. La constitution de 2011 le stipule dans son article 5 (*Cf.* Constitution du Maroc 2011).

La loi organique N°16. 26 parue dans le B.O. N°6816 datant du 26 septembre 2019 est organisée en 35 articles. (35 articles dont: les dispositions générales (2 articles); l'amazighe dans l'enseignement (5 articles); l'amazighe dans la législation et au parlement (3 articles); l'amazighe dans les médias (5 articles); l'amazighe dans la création littéraire et artistique (3 articles); l'amazighe dans l'administration publique (6 articles); l'amazighe dans l'environnement linguistique (3 articles); l'amazighe dans la justice (1 article); la procédure de suivi de la mise en œuvre de la loi organique (4 articles): la disposition finale (1 article).)

Après l'administration publique et les médias, l'enseignement est l'un des secteurs à quoi la loi organique a consacré 5 articles dans lesquelles elle stipule que « *l'amazighe est une langue pour tous les Marocains* » (Article 3, La loi organique N°16. 26, B.O. N°6816, 26 septembre 2019) et exige:

- L'intégration de l'amazighe d'une manière progressive à l'école publique et privée (Article 4);
- La généralisation de l'enseignement de l'amazighe au primaire, au secondaire collégial et qualifiant (Article 5);
- La création de filières et d'unités de recherche spécialisées dans la langue et culture amazighes dans l'enseignement supérieur (Article 6);
- L'intégration de l'amazighe dans les programmes d'alphabétisation et d'éducation non-formelle (Article 7);
- La prise en considération, dans l'élaboration des curricula et des programmes de l'amazighe, des différentes variantes de l'amazighe au Maroc (Article 8).

C'est en 2021 que le gouvernement s'est lancé dans le suivi de la mise en œuvre des dispositions de cette loi organique en créant une commission de suivi de ce projet avec la participation de différents secteurs (éducation et formation, culture, médias. . .). L'objectif est d'arrêter les modalités de mise en œuvre de l'officialisation de l'amazighe et de fixer un échéancier de réalisation des actions à entreprendre qui va de 01 an à 10 ans (*Cf.* Circulaire du Chef du gouvernement N°1234 relative à la mise en œuvre du plan gouvernemental intégrée pour la mise en œuvre du caractère officiel de la langue amazighe, 30 juin 2021).

La langue amazighe: de l'aménagement de son corpus à son enseignement

Les concepts de norme graphique et de production didactique, traités pour le cas corse, sont aussi valables pour l'amazighe du moment où il s'agit d'une part, d'une langue écrite, avant 2003, en 03 graphies (arabe, latine et tifinaghe) (Cf. El Barkani 2013) et d'autre part, une langue dont l'accès à l'école nécessite son équipement en outils didactiques et supports pédagogiques facilitant son enseignement-apprentissage. Quelle norme graphique pour l'amazighe ? et avec quelle production didactique, l'amazighe a-t-il pu intégrer l'école ?

Quelle norme graphique pour la langue amazighe?

Comme a été noté ci-dessus, l'aménagement du corpus est l'un des premiers objectifs de l'IRCAM depuis sa création d'où la nécessité de choisir une graphie avec laquelle la langue amazighe sera écrite et enseignée avant d'aller vers l'équipement de la langue en grammaire, conjugaison, lexique. . .

L'écriture est en usage au Maroc et en Afrique du Nord depuis l'Antiquité, la preuve en est un ensemble d'inscriptions qui ont été découvertes, gravées sur des stèles funéraires ou commémoratives (Boumalk 2002) depuis des siècles. Ainsi, écrire l'amazighe en libyque était une réalité qui subsiste encore avec les caractères tfinaghes. Mais, suite à un contact avec d'autres cultures et d'autres langues, la langue amazighe s'est écrite avec d'autres caractères donc, son rapport à l'écriture est devenu un rapport « externe » et la langue a baigné dans l'oralité.

Enfin le passage à l'écrit, qui n'implique pas forcément la mise à mort de l'Oralité (source d'identité), peut fournir aux sujets sociaux issus de la culture orale un meilleur moyen d'affirmation de soi, vu que l'écrit jouit d'une authenticité et d'une valorisation plus grandes que l'oral en raison des possibilités, non seulement, de contrôle qu'il offre, mais aussi de spécialisations scientifiques (...) Cohabitant en Afrique du Nord avec l'arabe (parlé et écrit) et le français, le berbère est noté en caractères arabes et latins, d'où son rapport le plus souvent « externe » avec l'écriture. (Cadi 1991: 92)

Ce rapport « externe » à l'écriture a fait que l'amazighe s'est écrit, depuis des siècles, avec l'alphabet arabe ou latin et seuls les Touaregs se sont servi et se servent encore de l'alphabet tfinaghe pour écrire leur langue maternelle qui est langue nationale au Niger et au Mali.

Pour choisir un système d'écriture pour la langue amazighe, trois groupes d'intellectuels défendaient leurs points de vue pour choisir l'une des trois graphies: les défenseurs de la graphie arabe exigeaient l'écriture de la langue amazighe en caractères arabes; les défenseurs de la graphie latine voyaient dans les caractères latins, un outil pour diffuser la langue amazighe à travers des écrits scientifiques en caractères latins; et les défenseurs de la graphie tfinaghe, eux, tentaient de conserver à la langue amazighe les caractères qui lui sont propres et qui symbolisent l'identité amazighe.

En janvier 2003 (sessions de travail du 30 et 31), le Conseil d'administration de l'IRCAM après consultation de plusieurs composantes (associations, partis politiques, universitaires, chercheurs...) a opté pour le tfinaghe comme graphie officielle de la langue amazighe. Une décision qui a pris effet après l'approbation de sa Majesté le Roi Mohammed VI, le 10 février 2003.

Une fois la question de la graphie tranchée, l'IRCAM s'est focalisé sur l'équipement de la langue en grammaire, dictionnaires... avant d'aller vers la production didactique en vue d'assurer l'introduction de l'amazighe dans l'école primaire en septembre 2003.

La production didactique et l'accès de l'amazighe à l'école

Le 1^{er} curriculum de la langue amazighe publié dans le livre blanc en 2002 vise le développement de la compétence communicative chez l'apprenant dans ses deux volets « compétence orale » et « compétence écrite » en prenant en considération l'âge et le degré du développement de l'apprenant. D'autres compétences sont visées comme les compétences stratégiques, les compétences culturelles (symboliques et encyclopédiques), les compétences méthodologiques et les compétences technologiques (cf. Note ministérielle N° 108, 2003). Pour développer ces compétences, le manuel de la langue amazighe, dans ses éditions 2003-2008 (depuis l'instauration de l'enseignement de l'amazighe en 1^{ère} année du primaire jusqu'à l'atteinte de la 6^{ème} année du primaire) s'est fixé comme objectif, d'assurer une certaine continuité dans l'accueil de l'enfant de son environnement familial à l'école, et ce, en prenant en considération la variante de l'élève dans les niveaux les plus bas (1^{ère} et 2^{ème} années) dans l'objectif de s'ouvrir sur les autres variantes dans les niveaux intermédiaires (3^{ème} et 4^{ème} années) et de se référer à un amazighe unifié dans les niveaux avancés (5^{ème} et 6^{ème} années) tout en prenant en compte les consignes et les structures linguistiques communes aux 3 variantes de l'amazighe et utilisées depuis la 1^{ère} année. Il s'agit d'une répartition en 03 cycles : un cycle de base, un cycle intermédiaire et un cycle avancé.

« 1. un cycle de base où l'enseignement-apprentissage de l'amazighe est fondé exclusivement sur les données géolectales selon les régions; le choix de l'enseignement dans les variantes géolectales a pour objectif d'assurer la continuité linguistique et culturelle entre le milieu familial et environnemental et l'espace scolaire;

2. un cycle intermédiaire enrichi où l'enseignement-apprentissage de l'amazighe est dispensé dans les géolectes selon les régions, mais avec une ouverture sur d'autres géolectes, le géolecte de base se trouve ainsi enrichi par la synonymie et la variation lexicale, morphosyntaxique et prosodique propres aux géolectes des autres régions. Cette démarche est censée contribuer à amorcer l'intercompréhension entre les parlers de l'amazighe;

3. un cycle avancé où l'amazighe est enseigné en tant que code polynormé enrichi par ses géolectes et développé en vue de préparer les utilisateurs à une fluidité et une perméabilité interdialectales et devenir performant dans des situations de communication à large spectre. » (Agnaou 2009: 24)

Ce curriculum a été révisé en 2021 (*Cf.* programmes et orientations pédagogiques du cycle primaire, MEN, juillet 2021; Décision du Ministre de l'éducation nationale, de la formation professionnelle, de la recherche scientifique N°080-21 datant du 29 juillet 2021 relative au nouveau curriculum du cycle primaire). Cependant, dans les nouveaux programmes de l'amazighe (édition 2021), les nouveaux manuels scolaires à mettre en œuvre en classe d'amazighe à compter de septembre 2021 doivent prendre en considération 02 cycles: 1^{er} cycle (1^{ère}-3^{ème} année) et 2nd cycle (4^{ème}-6^{ème} année). Dans le 1^{er} cycle il s'agit de privilégier la variante de l'apprenant en s'ouvrant en 3^{ème} année sur les expressions et le vocabulaire des autres variantes. Dans le 2nd cycle et en 4^{ème} année surtout, une ouverture sur les autres variantes en grammaire et conjugaison est exigée en vue de développer une intercompréhension interdialectale et préparer l'apprenant à une langue amazighe unifiée et enrichie par ses variantes en 5^{ème} et 6^{ème} année du primaire. La nouvelle approche adoptée dans la progressivité de la standardisation de l'amazighe dans les manuels scolaires reste à évaluer une fois les manuels scolaires élaborés et mis en œuvre en classe.

Au moment où le curriculum de la langue amazighe, paru en 2002 (Livre blanc) et révisé en 2021, donne les orientations relatives à l'enseignement de l'amazighe dans le primaire, des notes ministérielles organisent cet enseignement.

La révision des programmes de la langue amazighe (*Cf.* programmes et orientations pédagogiques du cycle primaire, MEN, juillet 2021; Décision du Ministre de l'Education nationale, de la formation professionnelle, de la recherche scientifique N°080-21 datant du 29 juillet 2021 relative au nouveau curriculum du cycle primaire.) s'inscrit dans la lignée de la révision des programmes du primaire telle que énoncée dans la loi-cadre N°17.51 relative au Système d'éducation et de formation et de la recherche scientifique parue dans le B.O. N°6805 datant du 19 août 2019 « Révision des approches, des curricula et des programmes pédagogiques (Préambule) ». Les nouveaux programmes de l'amazighe sont donc en complémentarité avec les programmes des langues arabe et française en termes de types de discours, de thèmes, de nombre de séquences didactiques. . . L'objectif est de créer des passerelles entre les langues enseignées au Maroc en vue d'en faciliter à l'apprenant ses apprentissages. Ces programmes constituent la base des concepteurs des manuels scolaires pour répondre aux cahiers de charges relatifs à l'élaboration des manuels scolaires de l'amazighe.

Le manuel scolaire de langue amazighe constitue l'un des supports didactiques de référence pour l'enseignement/apprentissage de la langue. Vu qu'il est question d'une langue en cours de standardisation et écrite officiellement en tifnaghe, le manuel doit représenter les choix linguistiques et didactiques de ses concepteurs en collaboration avec les chercheurs de l'IRCAM et le MEN. Ce sont les cahiers de charges relatifs à l'élaboration et la conception du manuel scolaire qui en définissent les contenus tandis que ce sont les notes ministérielles qui organisent l'enseignement de cette langue (ressources humaines mobilisées, formations exigées, volumes horaires alloués aux activités, emploi du temps. . .). Cependant, en présence d'une variation interdialectale et intradialectale de l'amazighe, quelle langue enseigner à l'école et comment ?

Quelle langue amazighe enseigner à l'école primaire ?

Si pour le corse, comme il a été noté précédemment, on a opté pour l'enseignement d'une langue polynomique, au Maroc par contre, le Curriculum de la langue amazighe dans sa première édition de 2002 ou dans sa version révisée en 2021 a noté qu'il faut, dans l'élaboration des manuels scolaires de l'amazighe, respecter certaines dispositions primordiales fixant la langue à enseigner. Il s'agit surtout de la nécessité de:

- Elaborer une langue amazighe unifiée et enrichie ayant comme référence les dialectes de l'amazighe;
- Mettre à la disposition de l'apprenant de l'amazighe un manuel scolaire, et adapter son vocabulaire aux spécificités régionales de la langue;
- Privilégier les structures linguistiques communes entre ces variantes dialectales et leur donner la priorité dans les manuels scolaires et les supports didactiques;
- Renvoyer, en cas d'absence d'un terme commun, à la néologie, au vocabulaire amazighe employé en arabe dialectal marocain, et dans les autres variantes amazighes en usage dans les régions amazighes (Algérie, Niger, ...).

Il s'agit donc, de l'enseignement d'une langue amazighe unifiée et enrichie par ses variantes.

Comment l'amazighe est enseigné en classe?

Depuis l'introduction de la langue amazighe à l'école primaire en 2003, un volume horaire de 03 heures hebdomadaires lui est consacré, au moins dans les textes organisant son enseignement-apprentissage (curriculum de 2002, notes ministérielles N°108, 130, 133...). Ce volume horaire qui a été maintenu dans les nouveaux programmes du primaire de 2021 a été souvent estimé par les acteurs pédagogiques et surtout les enseignants de l'amazighe comme insuffisant pour envisager la maîtrise de la langue amazighe par les apprenants en tant que langue officielle à l'instar de la langue arabe (Cf. El Barkani 2013; El Baghdadi 2016).

L'enseignement-apprentissage de l'amazighe s'organise en 06 activités, à raison de 15 à 30 ou 45 minutes par activité (Cf. Note ministérielle N°133). Il s'agit d'activités développant la compétence orale (Activités d'expression orale et de communication (ⵎⵓⵏⵉⵎⵉⵏⵉⵏⵉ)) et d'autres développant la compétence écrite (Activités de lecture (ⵜⴰⵎⴰⵣⵉⵔⵜ), d'écriture (ⵜⴰⵎⴰⵣⵉⵔⵜ) et de grammaire et conjugaison (ou activités relatives au fonctionnement de la langue) (ⵜⴰⵎⴰⵣⵉⵔⵜ ⵏ ⵜⴰⵎⴰⵣⵉⵔⵜ ⵏ ⵜⴰⵎⴰⵣⵉⵔⵜ)). Des activités ludiques (ⵎⵓⵏⵉⵎⵉⵏⵉ) sont programmées aussi à la fin de chaque semaine d'enseignement-apprentissage ainsi que des activités d'évaluation (ⵜⴰⵎⴰⵣⵉⵔⵜ ⵏ ⵜⴰⵎⴰⵣⵉⵔⵜ) des apprentissages et de soutien (ⵜⴰⵎⴰⵣⵉⵔⵜ).

Les anciens programmes de l'amazighe mis en œuvre en classe d'amazighe depuis 2003 organisent les manuels scolaires en 08 unités didactiques autour de

thématiques relevant de la vie quotidienne de l'apprenant et développant sa compétence orale et écrite en langue amazighe. Les nouveaux programmes qui sont mis en œuvre depuis septembre 2021 organiseront les manuels scolaires en 06 unités didactiques autour de thématiques relevant progressivement de l'environnement immédiat de l'apprenant jusqu'à son environnement national et universel (niveaux avancés: 5^{ème} et 6^{ème} année). Chaque unité d'enseignement-apprentissage s'achève par un projet à réaliser par l'apprenant en vue de mettre en œuvre ses acquis tout au long des semaines d'enseignement-apprentissage de l'unité en question.

Si les anciens programmes adoptent l'approche communicative, ces nouveaux programmes visent de plus, le développement d'une pédagogie différenciée permettant à l'enseignant de gérer des classes hétérogènes (en présence d'apprenants amazighophones et arabophones) et d'une pédagogie de projet incitant l'apprenant à réaliser des tâches dans le cadre d'une approche fonctionnelle-notionnelle.

Cette révision des manuels scolaires déclenchée en avril 2021, date de publication des nouveaux programmes de la langue amazighe, s'est généralisée en septembre 2022 sur les 6 années du cycle primaire.

L'enseignement-apprentissage de l'amazighe a suscité des enquêtes de terrain depuis 2003, nous prenons dans cet article comme exemples les 02 études (Études réalisées par les chercheurs: Mhamed EL BAGHDADI, Abdeslem KHALAFI, Fatima AGNAOU, Myriem DEMNATI, Bouchra EL BARKANI, Benaissa ICHOU, Mustapha SGHIR, L'Houssaine EL GHOLB, Kamal AQUA, Abdallah BOUZANDAG, en collaboration avec le MEN (AREFs, délégations, acteurs pédagogiques concernés par l'étude (inspecteurs, enseignants, ...)).) réalisées par les chercheurs du Centre de la recherche didactique et des programmes pédagogiques de l'IRCAM en 2010 et 2012.

Il s'agit dans la 1^{ère} étude de *l'évaluation du degré de maîtrise de la lecture et de l'écriture en tifinaghe* (2^{ème} année du primaire). (L'étude a été réalisée en 2010 au niveau de 6 Académies régionales d'éducation et de formation (AREFs): Souss-Massa-Draa; Taza-Al Hoceima -Taounate; Marrakech -Tansift -El Haouz; Doukkala -Abda; Fès- Boulmane; Rabat- Salé- Zemmour -Zair et sur un échantillon de 1100 apprenants. Ses résultats viennent d'être publiés en 2022.) Elle a montré, entre autres, que le tifinaghe ne pose aucun problème aux apprenants en écriture comme en lecture. Par contre, les élèves rencontrent des difficultés en production écrite.

La deuxième étude est *l'évaluation des apprentissages en lecture et production écrite en amazighe* (4^{ème} et 6^{ème} du primaire) (L'étude a été réalisée en 2012 au niveau de l'AREF Souss-Massa-Draa et sur un échantillon de 1597 apprenants.) (Cf: El Baghdadi 2016). L'étude a montré, entre autres, que les apprentissages en amazighe au niveau de la lecture et de la production écrite progressent en 4^{ème} et 6^{ème} mais des difficultés liées à la compréhension écrite ont été soulevées. Les stratégies de lecture ne sont pas développées chez les apprenants.

Suite à ces deux études, il s'est avéré que les conditions dans lesquels se fait l'enseignement-apprentissage de l'amazighe impactent les apprentissages des élèves. La discontinuité des apprentissages par exemple qui est liée surtout à la

non-généralisation de l'enseignement de l'amazighe sur le plan vertical, mène à un dysfonctionnement relativement au rapport entre le savoir enseigné en classe et le savoir assimilé par les apprenants. La formation des enseignants, le niveau scolaire des parents et la langue parlée en famille impactent le degré d'implication des élèves en classe et le développement de la compétence ciblée dans les programmes d'étude.

Pour améliorer l'enseignement-apprentissage de l'amazighe et suite aux résultats des études susmentionnées, nous estimons qu'il est nécessaire d'avoir les conditions favorables à cet enseignement. D'une part, il faut appliquer les instructions officielles et suivre les orientations pédagogiques fixées par le curriculum de l'amazighe, les notes ministérielles et les guides de l'enseignant. D'autre part, il faut, entre autres, veiller à la généralisation de l'enseignement de l'amazighe sur les plans vertical (introduire l'amazighe dans le secondaire) et horizontal (sur tout le territoire marocain); former les ressources humaines nécessaires; allouer un volume horaire plus important à l'apprentissage de la langue; introduire l'écrit en tifinaghe au préscolaire.

Le ministère de l'Education nationale, de l'enseignement supérieur, de la formation professionnelle et de la recherche scientifique a publié un communiqué de presse le 30 décembre 2020 où il a noté qu', en collaboration avec l'IRCAM, il veillera, en plus de la révision des programmes du primaire faite en avril 2021, à la révision des manuels scolaires de l'amazighe, voire l'élaboration de nouveaux manuels du primaire ainsi qu'à la préparation du curriculum de la langue amazighe au cycle secondaire collégial. Si les nouveaux programmes de l'amazighe publiés en avril 2021 se croisent avec les nouveaux programmes des langues arabe et française en termes d'unités didactiques, de thèmes et de compétences développées, ces langues à l'école doivent faire l'objet d'une enquête de terrain pour voir l'impact de ces nouveaux programmes sur les apprentissages des élèves en langues tout en fournissant toutes les conditions favorables à l'enseignement-apprentissage de l'amazighe au même pied d'égalité qu'en arabe et en français (volume horaire, généralisation, ressources humaines ...).

Conclusion

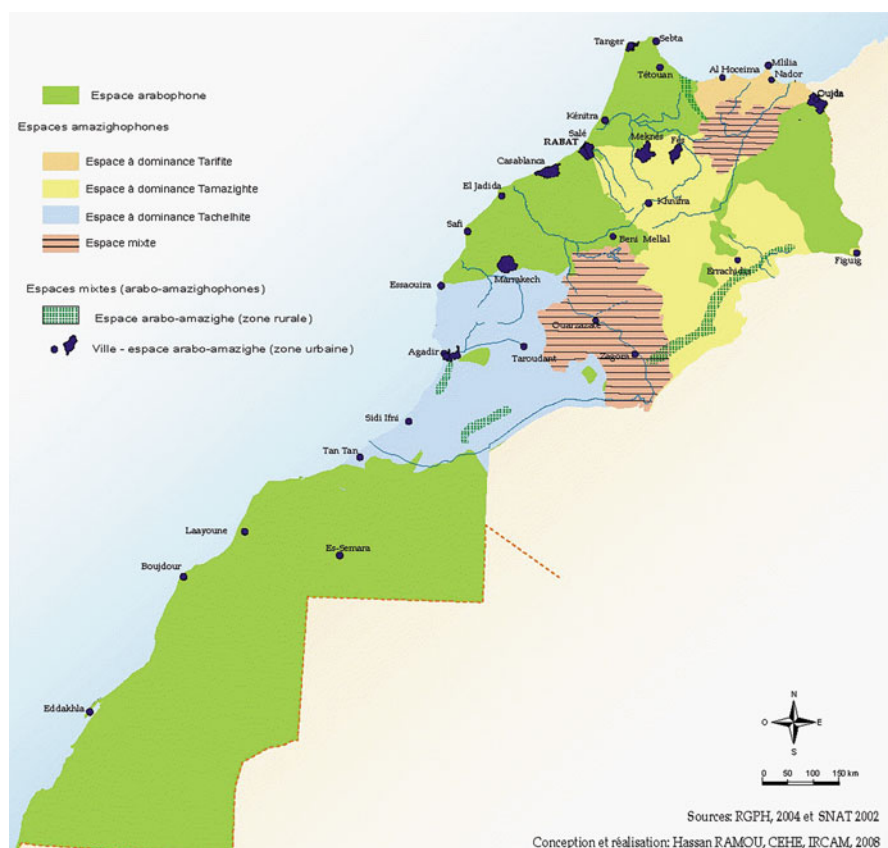
En revenant aux questions posées dans l'introduction de cet article interrogeant le rôle que joue l'école dans la mise en œuvre effective de l'officialisation de ces langues, d'après les données présentées dessus relativement à la situation de l'enseignement des langues corse et amazighe comme cas de langues minorées introduites à l'école, il s'est avéré que l'école, ne peut pas, seule, assurer la survie de ses langues en dehors de la famille, de la société et des différents domaines de la vie publique (média, administration. ...). Cependant, pour que la langue, ne reste pas enfermée à l'école, comme l'a souligné Jean-Marie Comiti (2005) pour le corse, il faut que la langue en question soit utilisée au quotidien, chose qui relève de la responsabilité de ses locuteurs.

Pratiquer une langue c'est maintenir sa vitalité, écrire une langue c'est en garder une trace et la sauvegarder. Tels sont les deux éléments primordiaux dans le maintien d'une langue et sa transmission intergénérationnelle. Si le maintien de la vitalité d'une langue et sa sauvegarde relèvent de la responsabilité des locuteurs de la langue, sa sauvegarde relève par contre de la responsabilité des locuteurs et des institutions.

Annexes



Carte 1 La Corse (Carte consultée le 9 août 2009 sur le Site Web: <http://www.tlfq.ulaval.ca/axl/europe/corsecarte.htm>)



Carte 2 Le paysage linguistique au Maroc (Carte élaborée par Hassan Ramou, CEHE, IRCAM, 2008)

Références bibliographiques

Documents officiels

Actes des 1^{ères} assises nationales des langues de France (2003, 4 octobre). Ministère de la culture et de la communication. Document consulté sur: http://www.culture.gouv.fr/dglf/politique-langue/assies/actes_assises.pdf

Dahir portant création de l'Institut Royal de la Culture Amazighe (IRCAM), 17 octobre 2001.

Document consulté sur: <http://www.ircam.ma/indexfr.asp?rd=351>

Décision N°080-21 du Ministre de l'Education Nationale, de la Formation Professionnelle, de la Recherche Scientifique relative au nouveau curriculum du cycle primaire (2021, 29 juillet).

Document consulté le 20 septembre 2021.

Décret N°74-33 du 16 janvier 1974. Document consulté sur: http://www.tlfq.ulaval.ca/axl/europe/France-loi_Deixonne-texte-1951.htm

- Discours royal de S.M. le Roi Mohamed VI à Khénifra, le 17 Octobre 2001. Document consulté sur: <http://www.ircam.ma/indexfr.asp?rd=382>
- La circulaire du Chef du gouvernement N° 1234 relative à la mise en œuvre du plan gouvernemental intégré pour la mise en œuvre du caractère officiel de la langue amazighe (2021, 30 juin). Document consulté le 10 mai 2022.
- La loi Deixonne (1951, 11 janvier). Document consulté sur sur: http://www.tlfq.ulaval.ca/axl/europe/France-loi_Deixonne-texte-1951.htm
- La loi organique N° 16. 26 relative à la mise en œuvre du caractère officiel de la langue amazighe (2019, 26 septembre). B.O. N° 6816. Document consulté sur: http://www.sgg.gov.ma/BO/bo_ar/2019/BO_6816_Ar.pdf
- Le Dahir n° 1-19-121 du 12 moharrem 1441 (12 septembre 2019) portant promulgation de la loi organique n° 26-16 fixant le processus de mise en œuvre du caractère officiel de l'amazighe, ainsi que les modalités de son intégration dans l'enseignement et dans les domaines prioritaires de la vie publique (2021, 1^{er} juillet). B.O. N° 7000. Document consulté sur: http://www.sgg.gov.ma/BO/bo_fr/2021/BO_7000_Fr.pdf
- Le livre blanc (2002, juin), Royaume du Maroc, Ministère e l'Education Nationale, de l'enseignement supérieur, de la formation des cadres et de la recherche scientifique. Commission de révision des programmes pédagogiques marocains de l'enseignement primaire, secondaire et qualifiant. Document consulté sur: http://www.men.gov.ma/dajc/Livre%20blanc/T2.prim/primamazigh_fichiers/editdata.mso
- Programmes et orientations pédagogiques du cycle primaire (2021, juillet). Ministère de l'Education Nationale, Rabat. Document consulté le 15 octobre 2021

Ouvrages et articles

- Agnaou F (2009) Vers une didactique de l'amazighe. Asinag. N° 2: Dossier: L'enseignement de l'amazighe: genèse, état des lieux et perspectives. Revue de l'Institut Royal de la Culture Amazighe, Rabat, pp 21–30
- Bentolila A, Germain B (2005) Apprendre à lire: choix des langues et choix des méthodes. UNESCO. Document consulté sur: <http://unesdoc.unesco.org/images/0014/001462/146229f.pdf>
- Blanchet Ph (2005) Minorations, minorisations, minorités: Essai de théorisation d'un processus complexe. In: Huck D, Blanchet Ph (Dir) Minorations, Minorisations, Minorités. Études exploratoires Presses Universitaires de Rennes [Cahiers de sociolinguistique N° 10]
- Boumalk A (2002) L'enseignement du berbère: de l'oralité à l'écriture: Tamazight face à son avenir. Publication de PASSERELLES-ONG, Thionville. Printemps-Eté [Passerelles. N° 24]
- Boyer H (2006) Présentation. In: ÉLA: Langues minorées, langues d'enseignement? Klincksieck, Didier Erudition, pp 261–263
- Cadi K (1991) Le passage à l'écrit: de l'identité culturelle à l'enjeu social. In Identité culturelle au Maghreb. Publication de la Faculté des Lettres et des Sciences Humaines de Rabat, Rabat
- Chiorboli J (2002) La codification des langues polynomiques: l'orthographe du corse. In: Caubet D, Sibille J (eds) Codification des langues de France. L'Harmattan, Paris, pp 141–153
- Comiti J-M (2005) La langue corse entre chien et loup. L'Harmattan, Paris
- Coyos JB (2004) Politique linguistique, langue basque et langue occitane du Béarn et du Gascogne. Elkarlanean S.L, Donostia
- Di Meglio A (1997) L'élaboration didactique d'une langue minorée: le corse. Problématique de l'enseignement du corse par l'approche sociolinguistique de documents didactiques entre 1974 et 1994. Thèse de doctorat (NR), Soutenue à l'Université de Corté
- Di Meglio A (2002) Emergence d'une langue et conditions de l'élaboration didactique d'un savoir: le cas du corse In La circulation des savoirs: la situation de l'enseignement de la langue et de la culture corse IUFM de Corse. Document consulté sur: <http://www.corse.iufm.fr/recherche/fichiers/savoirs.pdf>

- Di Meglio A, Comiti J-M (1999, décembre) La langue corse dans l'enseignement: norme(s) et fonctionnements. *Lidil*. N° 2 0: "Les langues régionales, enjeux sociolinguistiques et didactiques". Université Stendhal de Grenoble, Grenoble, pp 62-63
- Di Meglio A, Cortier C (2004) Le dépassement du conflit diglossique en corse: implications pédagogiques et didactiques chez les maîtres bilingues de l'école primaire. *Repères*. N° 29: Français et langues étrangères et régionales à l'école, Quelle interaction? INRP
- Di Meglio A, Cortier C (2005) La langue corse dans l'enseignement public. Trad. « Corsican Language in public Education ». Communication donnée au 10th International Conference on Minority Languages (Trieste, July 1st and 2nd 2005)
- El Baghdadi M (Sous la direction de) (2016) L'évaluation des apprentissages en lecture et production écrite en amazighe (4^{ème} et 6^{ème} du primaire) (en arabe) Rabat Publications de l'Institut Royal de la Culture Amazighe, Rabat
- El Barkani B (2013) Le choix de la graphie tfinaghe pour enseigner-apprendre l'amazighe au Maroc: conditions, représentation et pratiques, 2 vols. Thèse de doctorat en Didactique des Langues, Université Jean Monnet, Saint-Etienne
- Ettori F. (1975). L'enseignement de la langue corse. In J.-B. Marcellesi (ed.), *L'enseignement des langues régionales* (pp. 104-111). Paris.: Larousse [Langue Française N° 25].
- Fusina J (1984) Le corse. In: *Par les langues de France. Étude menée par le laboratoire de recherches interculturelles du CNRS. avec la collaboration technique de Rossana Vaccaro*. Centre national d'art et de culture Georges Pompidou
- Fusina J (2003) *Histoire de l'École en Corse*. Albiana éditeur, Ajaccio
- La Corse. Document consulté sur: http://www.tlfq.ulaval.ca/axl/monde/Plan_du_site.htm
- Lafont R, Ernest B (1982) *Langue dominante, langues dominées*. ediliG, Paris
- Lagarde Ch (2007) L'institution scolaire face à la diversité linguistique. In: Burban Ch, Lagarde Ch (travaux réunis par) *L'école, instrument de sauvegarde des langues menacées? Actes du colloque du 30 septembre et 1^{er} octobre 2005*. Université Perpignan Presses Universitaires de Perpignan, Perpignan, pp 11-24
- Marcellesi J-B (1986) Actualité du processus de naissance de langues en domaine roman. In: Marcellesi J-B (2003) *Sociolinguistique, Epistémologie, Langues régionales*. L'Harmattan, Paris
- Marcellesi J-B (2001) Parcours d'un sociolinguiste: de la langue corse au discours politique. In: Marcellesi J-B (2003) *Sociolinguistique, Epistémologie, Langues régionales*. L'Harmattan, Paris, pp 22-23
- Ottavi P (2007) L'enseignement d'un vernaculaire à l'école: repli identitaire ou ouverture à la modernité ? In: Burban Ch, Lagarde Ch (travaux réunis par) *L'école, instrument de sauvegarde des langues menacées? Actes du colloque du 30 septembre et 1^{er} octobre 2005* Université Perpignan. Presses Universitaires de Perpignan, Perpignan
- Renard R (2007) L'école, sûrement, mais pas seulement l'école. In: Burban Ch, Lagarde Ch (travaux réunis par) *L'école, instrument de sauvegarde des langues menacées ?*. Actes du colloque du 30 septembre et 1^{er} octobre 2005. Université Perpignan. Presses Universitaires de Perpignan, Perpignan
- Sibille J (2000) *Les langues régionales*. Flammarion, Paris
- Siguán M, Mackey WF (1986) *Education et bilinguisme*. Delachaux & Niestlé, Lausanne

Swadesh in the Sahara: Lexicostatistical Lists in Berber/Amazigh Languages and Their Loan Components

7

Anthony Grant

Contents

Introduction	153
Some Remarks on Sources	154
Contact Influences on Berber: A Synopsis	156
A Note on Early Semitic Influence: Cardinal Numerals	156
Arabic (and Other) Loans on Some Amazigh Swadesh Lists	157
Kabyle Berber/Taqbaylit (All from Arabic)	157
Other Berber Varieties	159
Comparison with Other High Borrowing Languages	161
Berber Loans in Northern Songhay Swadesh List Vocabulary	162
Conclusions	162
References	163

Abstract

This study examines the Swadesh lists for a number of Berber/Amazigh languages and demonstrates that the degree of borrowing (mostly from Arabic), while it is not as massive as in the vocabulary as a whole, is nonetheless very great (reaching over a third of the items in some varieties), and highlights the importance that Berber languages have for informing a solidly empirical basis for the elucidation of the possibilities of contact-induced linguistic change.

Introduction

The Berber languages have received a considerable amount of attention from historical linguists, though the majority of these have been dedicated Berberologists writing mainly for one another. Students of language contact have also paid a

A. Grant (✉)
Edge Hill University, Ormskirk, UK
e-mail: granta@edgehill.ac.uk

modicum of attention to them, though here again the pattern seems largely to be that of Berberologists speaking to one another away from more general audibility. That contacticians have much to learn from the examination of Amazigh languages is certainly true, and this brief chapter seeks to open this world up a little wider.

My point of departure is the content of the various lexicostatistical lists drawn up in the 1950s by Morris Swadesh (1909–1967). He documented the development of this work in a number of papers (especially Swadesh 1950, 1955). The lists are of various lengths in terms of the number of items: 100, 200, 207 composite, and 215. The 215-item list was original, the 100-item version being a pared-down form providing a list of words, which were supposed to have an especially high retention rate, while the 200-item list includes most of the 100 items, and the 207-item amalgamated the last two lists; occasional items were added subsequently. Altogether 224 items occur on the lists. The contents are a function of the areas where Swadesh did fieldwork or where he had acquainted himself with relevant languages: principally North America and Mexico, but also many in Europe and some spoken in Asia. Some temperate zone terms such as “to freeze,” “ice,” “snow” are not universally relevant. Versions of the list have nonetheless become indispensable items of equipment for investigators in a number of fields of linguistics, including descriptive, historical, and documentary linguistics.

It is an immense pity that Swadesh did not test his various wordlists against Berber data, because this would have sharpened his sense of what could and could not be borrowed between linguistic systems. Had he done so, his list would have been more nuanced. His reasons for not doing so are fairly clear when one considers the purposes behind his original investigations, which are solidly historical. Of the 11 case studies against which Swadesh tested his assumptions about retention rates for the purposes of glottochronology, the only Afro-Asiatic example involved Middle Egyptian as contrasted with Coptic. From this standpoint the omission of data from Berber languages is understandable; aside from Libyco-Berber inscriptions the first Berber linguistic data are from the Middle Ages, and what Swadesh was seeking from his studies was time-depth (Attic Greek to Dhimotiki being separated by about 2500 years, for instance, and Alfredian Old English from Modern English by a millennium or thereabouts).

Some Remarks on Sources

The sources for this study are disparate. A small number of Swadesh lists, using different versions of the list, are available in the public domain for a range of Amazigh varieties and no doubt others are in private hands. Of these, Thomas Penchoen’s 1973 list (Penchoen 1973) is most useful, for it includes all the items from both Swadesh’s and Cohen’s lists. It points out all the Arabisms (33 in the composite Swadesh list) and it also indicates which words of Arabic origin in the variety surveyed (Ayt Ndhir Tamazight of the Central Atlas) are expressed by other words (specifically those of Berber origin) in other varieties, and also which few words are not readily lexicalized in this variety.

Salem Chaker (Chaker 1984) used a modified 200-item list for Taqbaylit, Tashlhiyt, and Tuareg. This list included 36 items that were not found on any version of the Swadesh list, and loans from Arabic were noted. An important source of inspiration for the additions to this list was the work on Afroasiatic lexicon (then “Hamito-Semitic”) by the French Semitist David Cohen. Penchoen (1973) helpfully appends translations into Ayt Ndhir Tamazight of these words found in Cohen’s list. Chaker just as helpfully marked up Arabic loans with an asterisk in each list. For what it is worth, English Wikipedia has a Tarifiyt list taken from Kossmann (2009). It is a version of the combined 100- and 200-item lists. A similar list for Taqbaylit is found on Russian Wikipedia.

Souag (2013) is an invaluable exploration of what contact-induced linguistic change involves, and goes way beyond lexicon. Indeed, it says relatively little about lexicon, but it includes a version of the 100-item list for Siwi, with etymological information where available. Kossmann (2013) is a discussion of borrowing from Arabic into Northern Berber (that is to say, all the Amazigh languages apart from Zenaga, Tetserret, and Tuareg) and despite the title of the work it covers a much more extensive linguistic field, and does so with great attention to detail. We should also mention that Taine-Cheikh’s French-Zenaga dictionary (Taine-Cheikh 2010), though extensive, is in a sense an index to her longer work (Taine-Cheikh 2008) and she excludes Arabisms from this work, which other sources, even early ones such as Faidherbe (1877), list in their vocabularies.

It is not my intention to present full lexicostatistical lists for all or even a range of Berber languages; such a task is beyond my expertise. Instead, I am focusing on loan elements (especially from Arabic) in a range of lists. My findings for various Berber languages are taken from the following sources: Ayt Seghrouchen Zenati: Abdel-Massih (1971); Kabyle/Taqbaylit: Dallet (1985); Tamazight: Penchoen (1973), Tashlhiyt and Tuareg: Chaker 1984; Ghomara: Mourigh 2016; Siwi: Laoust 1932 (most forms) and Souag (n.d.) Not all these provide full sets of glosses for the Swadesh list items. In a comparative survey examining borrowed lexicon, Kossmann (2013) points out that the proportion of loans, mostly from Arabic, on the Swadesh 100-word list is highest in Ghomara, with 37%. I found 64 loans, all but two from Arabic (the exceptions are from Spanish), in the Ghomara equivalents of the longest form of the Swadesh list in the Ghomara-English vocabulary in Mourigh (2016: 449–529), and I cannot guarantee that all items on the composite form of the list are logged or glossed in that vocabulary. There will undoubtedly be more.

There have been several attempts to reconstruct elements of Proto-Berber phonology and lexicon, though no complete dictionary of Proto-Berber exists. Naït-Zerrad’s volumes (1998–2003) are excellent (though not perfect) as far as they go, though the coverage is incomplete as the work was not finished. Several attempts at listing non-Arabic elements in Berber languages have been made. One notes Newman (1882), a cataloguer of pre-Arabic roots in some Berber languages compiled by a Latinist and Arabist skeptic who also worked on Arabic and whose brother became a Roman Catholic saint, and Haddadou (2006), the latter (like Naït-Zerrad’s work) is the work of a berberophone compiler. Both these works draw elements from a range

of Berber languages (the choice is wider in the case of Haddadou, of course) and for each entry they make the varieties from which they take their data explicit.

Zenaga, the Tuareg varieties and Tetserrét constitute Southern Berber, which is a geographical rather than a true phyletic taxon. What unites them is that they are not Northern Berber (which can be construed as part of a phyletic taxon if Zenaga was the first Berber language to split from the group). A fuller study would interrogate Swadesh list material from other Berber varieties such as Awjila (van Putten 2013) as far as this is available or extant.

Contact Influences on Berber: A Synopsis

With respect to the body of data that I am examining, the following sources of influence are relevant, in chronological order: early Semitic (time unknown), Latin (first century BCE – c. 500 CE), Punic (exact dates uncertain but up to c. 500 CE), and Arabic (seventh century CE onwards). Since the nineteenth century, Spanish and French have also influenced some Berber languages over the past few hundred years. Influences from Latin (and via Latin, influence from Greek) and Punic are at best evanescent on the Swadesh lists but they are of more importance in the lexica of the languages as a whole and they should therefore be mentioned.

A Note on Early Semitic Influence: Cardinal Numerals

A set of numerals that is found in many Berber languages, including Zenaga (probably the first branch to split from the remainder of the Berber languages) derives from an as yet unidentified early Semitic source. We can regard these *faute de mieux* as Proto-Berber numerals. They are not taken from Punic, and they are not Arabic either – many Berber varieties have replaced many or most of their numerals with loans from Arabic. This is also the case with most or all Northern Songhay languages.

A note on transcription is needed. Vowel length is indicated by a colon, while the emphatic counterparts of *h t d s z r* are represented as *ħ ṭ ḍ ṣ ḥ ṛ*.

The reconstruction of the first decade of Proto-Berber numerals is from Prasse (1974) as listed in Blažek (1998). Emphatic consonants are expressed with the symbols mentioned above.

1. *yi:wa:n*
2. *si:n*
3. *karaḏ*
4. *hakku:z*
5. *sammu:s*
6. *saḏīs*
7. *(his)sa:h*

8. (*hit*)*ta:m*
9. *tiz(z)a:h*
10. *mara:w*

Of these, 5–9 have been assigned to an early Semitic loan stratum. There are doubtless other early Semitic loans in Berber languages but they have yet to be identified.

Arabic (and Other) Loans on Some Amazigh Swadesh Lists

I have used an aggregate of all Swadesh list items, more than 225 in number, on the lists below, including some cases where two glosses are provided on the English master list, such as “hold/take.” In such cases I have examined and interrogated the glosses separately.

Chaker’s list with the greatest number of loans is Kabyle/Taqbaylit, which is also the form of Amazigh, which I first encountered. Most of its loans are also found in other Northern Berber languages, less so in Southern Berber ones (Zenaga, Tetseret, Tuareg varieties; Southern Berber is a paraphyletic taxon, a set of languages, which do not constitute a historical grouping defined by positive exclusively shared innovations). I have therefore started with a listing of some Kabyle data, taken from Dallet (1985) and presented in a slightly modified transcription. There are of course many further Arabic loans, including some, such as Kabyle *ruḥ* “to go,” which happen to express concepts, which have not been included on the Swadesh lists.

Kabyle Berber/Taqbaylit (All from Arabic)

kull “all” (Siwi has *kullši* incorporating an Arabic emphatic deriving from Arabic *šayy* “thing”)

ḥayawa:n; lewḥuṣ: both mean “animal”

ʔla xaTer “because”

tšir; aḥruḥ “bird”

tənəffəs “to breathe”

dderya “children”

liba:s “clothing”

ḥsi:b “to count”

ʒazama “to cut”

ʃtəḥ “to dance”

ay^wbbar “dust”

tmanyā “eight”

bšid “far”

baba “father”

wəḥf “fear, desolation”

rifi “feather”
lhut “fish”
xamsa “five”
ʕarba “four”
fakya “fruit”
lhu “to be good”
ħfiʃ “grass”
semħess “to hear, listen”
ʃegged “to hunt”
mya “hundred”
if “to live”
bnadem “man, person”
nəzzəh “many”
yəmma “mother”
dayyiq “narrow”
təsʕa “nine”
ism “name”
lala “not”
ʒdid “new”
qdim “old” (Siwi *ʕarəf* is also a loan from Arabic)
faxs “person”
leʔeb “to play”
ʒbəd “to pull, to shoot” (Siwi *stət* “to pull” is also from Arabic)
lehwa “rain”
mədəwwəʔa “round”
ħakk “to rub”
məlħ “salt”
ʔməl “sand”
bħəʔa “sea”
wali “to see”
sebʕa “seven”
xəyyəʔa “to sew”
qʔii “sharp”
ʕanni “to sing”
riħa “to smell”
səʔta “six”
ʒəld “skin”
dəxxa:n “smoke”
ħdara; nəʔaq “to speak”
falaqa “to split”
naqar “to stab”
ʕukka:z “stick”
ʕadal “straight; right/correct”
ʕum “to swim”
qbadha “to take”

ʕaʃra “ten”
māʕaʃ; irqiq “thin”
xəmməṃ “to think”
tlata “three”
ḍegger “to throw”
šedd “to tie”
ttsedʒra “tree”
brəm; dewwer “to turn”
əʃrin “twenty”
ħmu “warm”
aʃ “what”; *maʔ* “what?” (AS)
lyaba “woods, forest”
xdəm “to work”

Other Berber Varieties

Almost all these forms, which are not found in our Kabyle materials, come from spoken Arabic varieties and are recorded in Siwi unless otherwise noted; AN – Ayt Ndhir Tamazight, AS: Ayt Seghrouchen Zenati; forms from Tarifiyt, which are not also marked as borrowings in these other Berber varieties or in Taqbaylit, are taken from Kossmann 2009; most of these are also recorded for Ghomara: Mourigh 2016; we note that some forms have the Arabic definite article metanalyzed as part of the Berber stem.

u “and” (Tarifiyt)
ləqʃur “bark”
aʃməl “bad”
qriš “to bite” (AS)
nəʃʃəs “to blow” (AS)
lʕil “child” (AN)
ləymam “cloud” (AS)
iyum “day”
mut “to die” (Siwi)
bħət “to dig”
wsəx “dirty” (AN)
ħfu “dull” (AN)
tabəʔawt “egg”
asmin “fat/grease”
rriš “feather” (AN)
ħəbba “a few”
šwin “to fight” (AN; Siwi *mqobət* is also a loan from Arabic but from a different stem)
lʕafit “fire” (SN)
simək “fish”
zzəl “to flow”

- twərdət* “flower”
nfər “to fly”
ʒəmməd “to freeze” (AN)
ʕamma: “full” (Tarifiyt)
mzyan “good” (AN)
ššʕəy “hair”
aTqil “heavy”
dəqq “to hit”
itħəl “inside”
ləbhərat “lake”
dħəs “to laugh” (Tarifiyt)
tawriqt “leaf” (AN)
ʕaʕrawi “left (hand)”
siqayl “leg”
skn “to live” (AS)
aTwil “long”
bəzzaf “many,” *kuma* “much”
žžbal “mountain” (AN)
qərrəb “near” (AN)
ʒʒiřət “night” (Tarifiyt; Siwi has Arabic-derived *əllilət*)
la “not”
uħra “other”
dfəʕ “to push” (AN)
ʕəħiħ “right, correct”
nnaħr “river”
lməhwəl “road”
ħarq “root”
lkkurda “rope” (AS, from Spanish or Italian), *tiSəmmət* (Siwa)
ʕml “rotten” (AS)
xrbs “to scratch” (AS)
rħabb “seed” (Tarifiyt)
qqim “to sit” (AS; *ʕanʕan* occurs in Siwi)
ajwwa “sky”
aħkik “small”
ttfuħ “to smell” (Tarifiyt)
ttəldž “snow”
ša “some” (AN)
ziyər “to squeeze” (AN)
piká “to stab” (Tarifiyt, from Spanish *picar*; note the atypically Berber /p/, which is fairly frequent in Tarifiyt and more so in Ghomara; Siwi has borrowed local Arabic *čuk*)
tyuyət “stone”
səggəm “straight” (AN)
amaʔa:bu:s “tail” (AS)

gǝa: “thick” (Tarifiyt; Ghomara has *gurdu* from Spanish, which takes Arabic inflection)

tawaḥ “to throw”

asen “tooth” (AS)

tišayra “tree”

ṯayən “two” (Tarifiyt)

ṣuqq “to vomit” (Tarifiyt)

ṛəšš “to wash”

wasəʕ “wide”

ayərbi “wind” (AN)

afir “wing”

msəḥ “to wipe” (Tarifiyt)

ssənt “year”

sfər “yellow”

axbaš “claw”

brā (AS), *ṣəmməʕ* (AN) both: “full”

An early Latin loan of note is *anzlus* “child” in Ahaggar Tuareg, which is ultimately from Latin *angelus* “angel” (perhaps borrowed ironically). But unless indicated otherwise the loans on the lists have a Semitic origin.

Comparison with Other High Borrowing Languages

An indication of how Tarifiyt, as an example of a high-borrowing Berber language, fares on the Swadesh list in terms of borrowings when it is compared to other heavy-borrowing languages can be seen in the table below, expanded from statistics in Haspelmath and Tadmor (2009; the Swadesh list calculations, expressed as percentages are my own). These are taken from a sample of 41 languages drawn from most parts of the world.

Language	Sample size	% loans overall	100-item	223-item
Selice Romani	1431	62.7	27	34.6
Tarifiyt Berber	1526	51.7	17	32
Gurindji	842	45.6	19	40
Romanian	2137	41.8	9.5	14
English	1504	41.8	9	15.5
Saramaccan	1089	38.3	28	39
Ceq Wong	862	37.0	10	27
Japanese	1975	34.9	17	18
Indonesian	1942	34.0	9	16
Bezhta	1344	31.8	2	7

(Gurindji is used in the Northern Territory of Australia; Ceq Wong is an Austroasiatic language of Malaysia; Saramaccan is a creole language of Surinam; and Bezhta is a Nakh-Dagestanian language of the eastern Caucasus.)

This score indicates the massive importance of borrowed items in Berber vocabulary (and also in phonology, through the transmission of borrowed phones such as /q/) even at the most basic level. It also suggests that the 100-item list is, as Swadesh predicted, the least permeable layer of a language's vocabulary with respect to contact influences, with the longer Swadesh list not too far behind in conservatism.

Berber Loans in Northern Songhay Swadesh List Vocabulary

Berber languages have also been donors of items of basic vocabulary. Tasawaq, Tadakshak, Tagdal (and Tabarog), and Korandje (sometimes known as Kwarandzyey) are all languages spoken in the Sahel (at the oasis of Tabelbala, Algeria in the case of Korandje, in the Azawagh Valley in Niger in the case of the others, where Emghedessie, the now extinct languages of Agadez also belonged to this group). In these languages, which have been put together as a Northern Songhay group (although data to link them as a hession in themselves in the form of exclusively shared innovations is limited), the elements of basic Songhay morphology and a few hundred items of Songhay lexicon are expanded by hundreds of borrowings from Berber languages (often but by no means only Tuareg ones: Souag 2015 points out the crucial role of a language similar to Tetserrét on Tadakshak), with further loans coming into these languages from neighboring African languages (such as Fula and Dioula), Arabic, and latterly French. The proportion especially of Berber loans on the Swadesh lists of these languages is considerable (although data for a fuller study are not readily available) and it includes the names of common body parts, for instance, *i-lis* “tongue,” in addition to forms found in Berber languages, which are taken from Arabic, such as *fikr* “to think.” be Benítez-Torres and Grant (2017) discuss these, especially those spoken in the Azawagh Valley, and contains references to much literature on the varieties.

Tadakshak provides a fine case study of this. Christiansen-Bolli (2010) includes a list of 290 items, which Tadakshak shares with Songhay languages; these include some loans from French and Arabic, however, which non-Northern Songhay varieties use. Ninety-nine of these items appear on the 100- or 200-item Swadesh lists for Tadakshak, for which we have 177 glosses available in the Christiansen-Bolli material. At least 20% of the non-Songhay items are from a Berber language. Some examples from the description of Tadakshak by Christiansen-Bolli (2010) are *yiltág* “to be heavy,” *yidder* “to hold,” *a-múudər* “animal,” and *a-yázar* “lake, pond.”

Conclusions

More than half the items on the composite Swadesh list are represented by words of non-Berber origin (overwhelmingly from Arabic) in at least one (and often more) of the Berber languages discussed in this chapter. The proportion of such loans

increases from the 100-item Swadesh list, via the 200-item and 215-item lists, to the longer wordlists, such as the 1460-item list, which underlies Kossmann's Tarifiyt record. Berber languages belong to a subset of languages (including Brahui, Chamorro, Garifuna, and Romani varieties), which have absorbed very high proportions of elements and this can be seen within even limited bodies of data (such as the composite Swadesh list). This highlights the importance and potential that Berber languages have for informing a solidly empirical basis for the elucidation of the possibilities of contact-induced linguistic change.

Acknowledgments The impact of Maarten Kossmann and Lameen Souag on this paper are immense; they are, however, not responsible for any claims made in this paper or for any (mis) use made by me of materials to which they have directed me. I am extremely grateful to Farid Ait Si Selmi, who introduced me to Taqbaylit during my undergraduate phonetic transcription classes in the Department of Language and Linguistic Science, University of York, in 1981.

References

- Abdel-Massih E (1971) A computerized lexicon of Tamazight: Ayt Seghrouchen dialect. University of Michigan, Ann Arbor
- Benítez T, Carlos M, Grant AP (2017) On the origin of some Northern Songhay mixed languages. *J Pidgin Creole Lang* 32:263–303
- Blažek V (1998) Berber numerals. *Arch orient* 68:149–168
- Chaker S (1984) Les emprunts arabes. In: Chaker S (ed) *Etudes de linguistique berbère*. Editions du CNRS, Paris, pp 216–229
- Christiansen-Bolli R (2010) A grammar of Tadaksahak. Rüdiger Köppe, Köln
- Dallet JM (1985) *Dictionnaire français-kabyle*. Peeters, Louvain
- Faidherbe L (1877) *Le zénaga des tribus sénégalaises*. Leroux, Paris
- Haddadou MA (2006) *Dictionnaire des racines berbères communes*. HCA, Algiers
- Haspelmath M, Tadmor U (eds) (2009) *Loanwords in the world's languages: a comparative handbook*. Mouton, Berlin
- Kossmann MG (2009) Loanwords in Tarifiyt, a Berber language of Morocco. In: Haspelmath M, Tadmor U (eds) *Loanwords in the world's languages. A comparative handbook*. Mouton, Berlin, pp 191–214
- Kossmann MG (2013) *The Arabic influence on Northern Berber*. Brill, Leiden
- Mourigh K (2016) A grammar of Ghomara Berber. Rüdiger Köppe, Köln
- Nait-Zerrad K (1998–2002) *Dictionnaire des racines berbères (formes attestées)*. Peeters, Paris/Louvain
- Newman F (1882) *Libyan vocabulary*. Trübner, London
- Penchoen T (1973) *Tamazight of the Ayt Ndhir*. University of California Press, Los Angeles
- Prasse K-G (1974) *Grammaire touarègue (dialecte de ahaggar) IV–V Nom*. Akademisk Forlag, Copenhagen
- Souag L (2013) *Berber and Arabic in Siwa (Egypt): a study in linguistic contact*. Rüdiger Köppe, Köln
- Souag L (2015) Non-Tuareg Berber and the genesis of nomadic northern Songhay. *J Afr Lang Ling* 36:121–143
- Souag L (n.d.) *Siwi dictionary*. Manuscript in compiler's possession
- Swadesh M (1950) Salish internal relationships. *Int J Am Ling* 16:157–167
- Swadesh M (1955) Towards greater accuracy in lexicostatistic dating. *Int J Am Ling* 21:121–137

-
- Taine-Cheikh C (2008) Dictionnaire zénaga-français (berbere de Mauritanie). Rüdiger Köppe, Köln
- Taine-Cheikh C (2010) Dictionnaire français-zénaga (berbere de Mauritanie): avec renvois au classement par racines du dictionnaire zénaga-français. Rüdiger Köppe, Köln
- van Putten M (2013) A grammar of Awjila Berber (Libya): based on Umberto Paradisi's material. Doctoral dissertation, Universiteit van Leiden

Auditory Word Recognition in Tashlhit/ Berber: Supraliminal Speech Priming

8

Fatima El Hamdi

Contents

Introduction	166
Tashlhit	167
Word Recognition Process (Speech Supraliminal Priming)	169
Frequency Pretest	171
Experiment 1: Form and Meaning in the Tashlhit Mental Lexicon	171
Method	172
Results	175
Discussion	177
Experiment 2: Morphological and Phonological Priming Effect in Tashlhit	178
Method	180
Results	181
Discussion	182
General Discussion	183
Conclusion	187
Appendix A (Stimuli Used in Experiment 1)	187
Appendix B (Real Words Used in Experiment 2)	189
References	189

Abstract

The present study investigates lexical access in Tashlhit, an understudied Hamitic-Semitic language in the psycholinguistic field. We tested the hypothesis that the root morpheme has a lexical status in the Tashlhit lexicon. With this end in mind, we also examined the semantic and phonological factors in lexical access. We report here on two lexical decision experiments using supraliminal priming technique in which the participants are consciously aware of the prime. We used the auditory modality where target and prime words were audible. The results showed a priming effect by pairs sharing the root morpheme with an

F. El Hamdi (✉)

Department of English, Mohammed V University in Rabat, Rabat, Morocco

e-mail: f.elhamdi@um5r.ac.ma

opaque or transparent semantic relation, but no priming effect was obtained by pairs sharing only phonological properties.

Keywords

Auditory priming · Amazigh/Berber · Word recognition · Root · Lexicon organization · Phonology · Semantics

Introduction

The role of morphological theories in understanding how the lexicon is organized has been the topic of research of a number of studies. This topic brings two contrasting views to the fore.

Some researchers claim that the lexicon consists of independent polymorphemic words. This approach has been advocated in Amazigh and other Semitic languages like Hebrew and Arabic, pointing to the fact that the morphology of the language is better explained using a word-based approach. Arguments have been presented from Hebrew verb morphology (Bat-El 1994; Aronoff 1994; Ussishkin 1999), and from Arabic verbal and nominal morphology (Hammond 1988; McCarthy and Prince 1990; Guerssel and Lowenstamm 1996; Benmamoun 2003).

Contrastively, others argue that polymorphemic words are decomposed into morphemes among which we cite the base form or the root (Cantineau 1950; McCarthy 1981; Galand 1984; Chaker 1990; Tobin 1990; among others). Evidence for the root-based theory has been provided from language games (Arabic: McCarthy 1981; Tashlhit: Lahrouchi 2004, 2018), metathesis (Prunet et al. 2000), and behavioral studies (Deutsch et al. 1997, 1998, 2000a; Boudelaa and Marslen-Wilson 2001, 2004a, b, 2005; Ussishkin and Twist 2009; Ussishkin et al. 2015). According to this view, the root is accessed very quickly in studies of language processing (In the present work, we adopt the definition of the root as the base form, which cannot be further analyzed, or decomposed morphologically).

The present study aims to contribute to the debate on the two views on morphological theory and discuss the theoretical implications for the organization of the lexicon. It investigates the notion of roots in Amazigh; more particularly in Tashlhit and we will attempt to answer the main research question of whether the root is a significant morphological unit in the Tashlhit lexicon. We are more intrigued in the role played by morphology in language processing. However, semantic, and phonological factors will also be examined throughout this chapter.

With this end in view, this research will discuss the significance of the root from an experimental perspective. The data analyzed comes, essentially, from the variety spoken in Ighrem N'Ougdal area and its surroundings: a place largely representative of the Tashlhit language. We will discuss data resulted from priming experiments, based on measuring the reaction times of the participants. The experiments aim to

test whether roots in Tashlhit have some psycholinguistic reality and, hence, whether they have any significant implication for the organization of the Tashlhit lexicon. This type of experiment generally exposes participants to a stimulus (prime), which influences their response to a subsequent stimulus (target).

Tashlhit

Amazigh is one of the branches of Hamito-Semitic/Afroasiatic language family. The term Tamazight [tamaziɣt] is the Amazigh word used to refer to the language in Morocco. Geographically speaking, the language covers “all of North Africa, the Sahara, and a part of the West African Sahel” (Chaker 2008) (For more details on the Berber dialects and the corresponding areas where these dialects are spoken, see Chaker (2008)).

Native speakers of the language in Morocco comprise 26.7% of the population, following the 2014 Moroccan census (The reported results about native speakers of Amazigh in Morocco have been subject to debates among activists). In Morocco, Amazigh is not restricted to only one variety but rather refers to three major ones: Tarifit in the north, Tamazight in the center and the southeast, and Tashlhit in the southwest and the High Atlas (El Mountassir 2003; Ameur et al. 2004). The three main varieties are “united” by the formal Standard Amazigh (Standard Amazigh has been introduced by IRCAM (L’Institut Royal de la Culture Amazighe) and has been taught at some primary schools in Morocco since 2003, before Amazigh has become the official language of Morocco, in addition to Arabic, in 2011). Tashlhit is acquired as a first language at home, and it has many dialects. However, it is worth mentioning that the dialectal variation we notice in Tashlhit is not, necessarily, due to geographic distance: it can be observed even in near areas.

The variety of focus in this analysis is spoken in the region of *Ighrem N’Ougdal*, which covers a number of towns among which we cite *Agwim* and the nearby tribes: *Tamstinte* and *Tidili* (*Agwim/Agouim* is a village in Sous-Massa-Drâa, situated between Tizi-n-Tichka, and the city of Ouarzazate. It is 70 km away from Ouarzazate.). *Tamstinte* consists of many small towns: *Tichekiouine*, *Aslen*, *Tadouyakheth*, and others; while towns in *Tidili* are *Ighourassen*, *Tigheramt*, *Igheroud*, and others.

The consonantal and vocalic inventories below do not differ from those of other Tashlhit varieties (We should note that Tashlhit is an oral language and has no standard orthographic system. The only writing system Morocco has for Amazigh is Tifinaghe-IRCAM, to write and read Standard Amazigh. Many points are raised for discussion on this matter, both from a sociolinguistic, and a language-policy point of view, but we will not get into more detail on this issue, for it is beyond the scope of the present study). The vocalic system of Tashlhit is simple, consisting of three main vowels (Fig. 8.1) (Basset 1952; Boukous 1987, 2009; El Moujahid 1979;

Fig. 8.1 Tashlhit vocalic system

/i/ /u/

/a/

Table 8.1 Tashlhit Consonantal System

Labials	Dentals	Palatoalveolars	Velars	Uvulars	Pharyngeal	Laryngeals
	t t ^ɕ		k k ^w	q q ^w		
b	d d ^ɕ		g g ^w			
m	n					
f	s s ^ɕ	ʃ ʃ ^ɕ		x x ^w	ħ	
	z z ^ɕ	ʒ ʒ ^ɕ		ʁ ʁ ^w	ʕ	ʕ
w	l	r r ^ɕ	j			

Bensoukas 2001) (It is important to highlight an important point about the vocalic system of the language. Some studies argue that the schwa is an epenthetic vowel in Tashlhit inasmuch as it is the phonetic realization of syllable nuclei (Coleman 2000). The opposing view is that the language is marked for its syllabic consonants, overriding the need of epenthetic vowels, since consonants can also occupy the syllable nuclei position (Dell and Elmedlaoui 1985, 1988).).

In contrast, the consonantal system is much more complex given that some consonants have their emphatic counterparts and others labialized counterparts (Table 8.1). In addition, each consonant has a geminate counterpart except for the laryngeal ʕ (We used an IPA system for all the symbols). Some consonants may have more than one geminate counterpart, and each form is attested in some particular area. For instance, the geminate correspondents of the labial w, the uvular ʁ, and the dental d^ɕ, may be realized as ww (*tawwuri* “work”) or gg^w (*lgg^wr* “escape, imperf.”), ʁʁ (*ʁʁ^wr^ɕ* “lay, imperf.”), or qq (*r^ɕqqa* “warm, imperf.”); and dd^ɕ (*aydd^ɕar^ɕ* “cheater”) or tt^ɕ (*ntt^ɕu* “jump, imperf.”), respectively.

Given that a stem is what is left when agreement markers are taken from the verbal form, we assume that Tashlhit verbs have four main stems: the aorist or imperative, the imperfective, the perfective, and the negative perfective (Dell and Elmedlaoui 1988, 1992, 2013; Iazzi 1991; Bensoukas 2001; El Mountassir 2003; Boumalk 2004; Laabdellaoui et al. 2012). The aorist, and the imperative, can be defined as an order, an intention, a wish, or a future. The imperfective refers to a completed action, and the perfective is used to express a current, continuous, or repetitive action (Bensoukas 2018).

On the other hand, Tashlhit nominal morphology is known for its derivational, and inflectional nouns. For the former, the two common categories are action and agentive nouns. Place (*amzday* “cemetery”) and instrument nouns (*amkraz* “tool used to plow”) are not productive in the language. For instance, for the root √g^wnu, we have *tigni* as action noun, *imgni* as agentive noun, and *issgni* as instrument noun

(Bensoukas 2014).¹ For inflectional nouns Tashlhit is characterized by three categories: gender, number, and state. Singular nouns are marked with an initial vowel, mostly *a-* (*argaz* “man”). Other nouns are marked with initial *i-* and *u-* (*ilm* “skin,” *udm* “face”). Plural forms are characterized by their concatenative, and non-concatenative morphology (Jebbour 1988; Idrissi 2001; Bensoukas 2016, 2018). They can be divided into sound (external/concatenative), broken (internal/non-concatenative) (e.g., *agadir* (sg.) – *igidar* “walls,” *aqwzi* – *iqwza* “holes,” *aglzim* – *iglzam* “poleaxes”), and mixed (combination of both external, and internal morphology) (e.g., *asrdun* – *isrdan* “mules,” *adʿadʿ* – *idʿudʿan* “fingers”) (Bensoukas 2016). Sound plurals are the most common in Tashlhit. They are marked with the initial vowel *i-* accompanied with the suffix *-n* (E.g. *ayarʿasʿ* – *iyarʿasʿn* “roads,” *argaz* – *irgazan* “men”). Feminine nouns are marked with the affix *t—t* (*tamyarʿt* “woman,” *tajjist* “mare (horse)”). Furthermore, action nouns are also inflected for state. They surface, either in their free state, or construct state (*arʿzzʿum* (free state) – *urʿzzʿum* (construct state) “opening,” *tamgra* (free state) – *tmgra* (construct state) “harvest”).

Word Recognition Process (Speech Supraliminal Priming)

In studies of language processing and word recognition, in particular, the significant role of morphology has been demonstrated using different modalities: the visual (Boudelaa and Marslen-Wilson 2005, 2013; Frost et al. 2005), the auditory (Emmorey 1989; Boudelaa and Marslen-Wilson 2013; Schluter 2013; Gwilliams and Marantz 2015), and the cross-modal priming. The main objective of these works was to test whether morphological structure is accessed during language processing.

The strong effect played by roots in the word recognition process has already been demonstrated in Semitic languages and Indo-European languages. Several priming studies revealed the facilitatory role played by root morphemes in lexical access, in early and late word recognition process in Arabic (Boudelaa and Marslen-Wilson 2001, 2005, 2013; Schluter 2013; Gwilliams and Marantz 2015; Al Kaabi 2015), Hebrew (Deutsch et al. 1997), English (Emmorey 1989; Bentin and Feldman

¹ Another category of nouns that has received special interest is referred to in the literature as *bu-nouns* (Bensoukas 2015a, 2015b).

- bu. + *agajju* *buwgajju* “strong-headed person”
- bu. + *ayyu* *buwyyu* “the one who sells butter-milk”
- bu. + *zzriʿfa* *buzzriʿfa* “seller of nuts/dried fruit”
- bu. + *lassʿall* *bulassʿall* “owner of the gymnasium”

The *bu* of *bu-nouns* might be attached to borrowings or native words. Bensoukas 2015a, b defines this affix as one that “expresses the generic notion of ‘the one with X’ where X stands for any noun.” In these studies, the author has given these nouns a fully comprehensive morpho-syntactic treatment. In addition, Bensoukas (2015a, b) also points to the similar behavior of *bab-* and *bu-*. Both affixes express ownership and are attached to nouns through periphrasis and affixation, respectively. For an exhaustive treatment of *bu-nouns*, see Bensoukas (2015a, b).

1990; Rastle et al. 2000), German (Smolka et al. 2009), and French (Longtin et al. 2003; Longtin and Meunier 2005). Yet, a complete theory on word recognition process calls for a consideration of the mental lexicon of different languages. Hence, cross-linguistic research is deemed necessary for that matter.

Supraliminal priming, or non-masked priming, aims to study the late stages of language processing (A set of pilot studies were conducted using speech subliminal priming/auditory masked priming with different compression techniques to detect identity priming. We used Kouider and Dupoux's (2005) technique, which obtained identity priming with French-speaking participants at 35%, and 40%, compression rates. In the pilot trials, the prime stimuli were compressed at 35% rate and compressed at fixed prime durations (240 ms, 260 ms, and 280 ms) following the technique used by Schluter (2013) with Moroccan Arabic data. The participants were asked, by the end of the experiment, if they had noticed the presence of any prime, and none of them had reported a positive answer. These priming techniques were tested in a total of six experiments, but none of them yielded a repetition priming effect). In supraliminal priming tests, the participant gets exposed to a prime stimulus, auditorily, or visually, at a long exposure duration, and it, either immediately precedes the target, or it precedes other interveners – then the target follows. The main idea of supraliminal priming is the conscious exposure of the participant to the prime stimulus for processing. The consciousness of the presence of prime stimuli is used as a strategy to examine lexical processing and integrative processes across words (Milin et al. 2017). In this type of experiment, only the prime and target are used, i.e., no reversed words, no masks, and no compression are relevant to supraliminal priming.

Priming tests have been used in different modalities. In the visual modality (Forster and Davis 1984; Frost and Bentin 1992; Becker 1980; Deutsch et al. 1997, 1998, 2000a, 2000b; Kunde 2004; Boudelaa and Marslen-Wilson 2005), the prime and target stimuli are visually exposed to the participant, unlike in the auditory modality where the prime and target are auditorily exposed (Blumstein et al. 1982; Emmorey 1989; Boudelaa and Marslen-Wilson 2004a, 2013; Balling and Baayen 2008; Dupoux and Mehler 1990; Ussishkin and Twist 2009; Gwilliams and Marantz 2015; Gwilliams et al. 2015; Ussishkin et al. 2015). Both visual and auditory exposure of the stimuli can be used, simultaneously, in an experiment and this is what is referred to as cross-modal priming (Marslen-Wilson et al. 1994; Boudelaa and Marslen-Wilson 2001, 2004b, 2015). In the present study, the auditory modality is used to examine aspects of word recognition in language processing, in the Tashlhit mental lexicon.

One vital advantage worth noticing about the auditory priming is that it gives a chance to preliterate children and illiterate adults to participate (Kouider and Dupoux 2005), especially when the language being investigated is spoken by many illiterate participants. From a linguistic point of view, testing a language in a spoken mode is more natural than testing a language in a visual mode; the spoken mode is the first – naturally acquired by children (Schluter 2013). Besides that, it is not always practical to present visual orthography of a language, which has no writing system (Kouider and Dupoux 2005; Schluter 2013). It is true that Arabic or Latin scripts can be used to present the orthographic structure of a word since there are participants, who use such system to write the language in various situations, e.g., as in text messages. However, participants may use different letters of the alphabet, or a construction of

two letters, to represent the same phoneme. For instance, we have both /r/ and /gh/ to present the phoneme /r/ in Moroccan languages. Numbers are also used as one way of writing some of the phonemes. For instance, 3, 4, and 5 are used to refer to /ʕ/, /ɣ/, and /x/, respectively. Moreover, using this informal system of writing excludes illiterate participants (mainly adults) who are not to be excluded from the sample population. Other than that, in studies of the consonantal root in Hebrew and Arabic, the orthographic representation favors consonants over vowels, which may result in an overlapping with the consonantal root. Hence, the visual modality does not escape from the bias by the visual orthography (Ussishkin et al. 2015).

In studies that tested semantic, along with morphological effect, results showed an inhibitory effect of semantic priming. In semitic languages, particularly in Moroccan Arabic (Schluter 2013), Hebrew (Deutsch et al. 1997), and Maltese (Ussishkin et al. 2015), root priming showed a facilitatory priming effect whereas semantics showed no priming effect, suggesting that semantics plays no role in lexical access facilitation (Deutsch et al. 1997). These results were attested using different modalities (visual lexical access: Boudelaa and Marslen-Wilson 2005, 2013; auditory priming: Emmorey 1989).

In the present chapter, experiment 1 uses a lexical decision task to contribute to this research area using new data to test whether there is root and semantic priming effect in Tashlhit. The goal is to examine at which point the stored lexical representations become available to language processing, assuming that the main objective of lexical processing is to make the stored information of a given word available (Frauenfelder and Komisarjevsky Tyler 1987).

Frequency Pretest

In word processing experiments, the lexical retrieval for words might be influenced by factors like frequency. Hence, we tried to estimate the occurrence frequency of the Tashlhit words we used in our priming experiments. We collected Tashlhit words from dictionaries, data in articles, and dissertations about Tashlhit, and also from fieldwork data collection (Lasri 1991; Iazzi 1991; Bensoukas 2014; Lahrouchi 2010; Boumalk 2004; El Mountassir 2003). Twenty subjects who haven't participated in either of the priming experiments in this study, and whose ages ranged from 18 to 45, were asked to rate a list of words based on how frequently they encountered a word on a 7-point scale (1 = never, 2 = once a year, 3 = once a month, 4 = once a week, 5 = once every 2 days, 6 = once a day, 7 = several times a day) (Balota et al. 2001; El Hamdi 2018). The list of words used in the survey were auditorily presented to the participants. In this research, we used all the stimuli that were rated at least 3 by 90% of the participants.

Experiment 1: Form and Meaning in the Tashlhit Mental Lexicon

In experiment 1, we tried to investigate the interaction between morphology and semantics in the mental lexicon of Tashlhit. We attempted to determine whether the root in Tashlhit has a lexical role, and whether its facilitatory effect, if any, depends on semantic relatedness. An examination of semantic and morphology interface is

necessary to determine what activates lexical access. Previous studies reported a facilitatory priming effect when the prime and target share morphology, but not semantics (Boudelaa and Marslen-Wilson 2005; Emmorey 1989). By contrast, others reported that morphological effect is dependent on semantic relatedness (Gonnerman et al. 2007; Kielar and Joanisse 2011). In this experiment, we used a lexical decision task in a supraliminal speech priming test, in which the participant is aware of the auditory heard prime. We examined whether it is the lexical representation of morphology, or that of semantics that becomes available to language processing in late stages of spoken word recognition in Tashlhit.

Method

This section reports on details about the participants who conducted the experiments. It also describes the experimental stimuli used in this experiment, along with the five priming conditions used to test the effect of, both morphology and semantics.

Participants

Fifty-six subjects participated in experiment 1 (22 males and 34 females). The age of the participants ranged from 18 to 42 with the mean age of 26 and the median age of 24. Two participants were removed for hearing problems, and 24 were removed for their accuracy rate having been lower than 75: only data points from 30 participants were considered in the analysis. We also trimmed away incorrect responses from the analysis. No subject participated in more than one experiment. All the participants were from *Ighrem Nougda* and *Tidili*, towns where all speak Tashlhit as their mother tongue (Three recruiters have participated in conducting the experiments. They tried to recruit participants from their neighbors, friends, friends of their friends, relatives, relatives of their friends, etc. Asking participants to take part in this study, and run the experiment on a laptop, was not an easy task. Hence, the recruiters played a tremendous role in collecting the data; a familiar person had far better chances of recruiting more participants than one who was stranger. The recruiters had to go from one “douar” to another within the same area to solicit more participants. An exhaustive description of the sample population used in this study is provided in El Hamdi (2018)).

Materials

Each participant heard a total of 156 target words with both real words and non-words, which were paired in five priming conditions:

1. Identity in which the prime and target were the same.
2. Morphology and semantics (+root +sem) in which the prime and target shared both the root and semantic features.
3. Morphology, but not semantics (+root –sem), in which the prime and target shared the root, but no semantic feature

4. Semantics but not morphology (–root +sem) in which the prime and target shared semantic features but no root.

Unrelated condition in which the prime and target were completely unrelated (Under the semantics condition, we used stimuli, which shared semantic features, rather than synonyms or semantic associates. We run a semantic pretest to select the words that share many semantic features. We first tried to produce as many words as possible, which share semantic features, and then we selected the pairs of words sharing the most number of the semantic features (El Hamdi 2018).). Based on frequency and semantic pretests, 16 target words were tested under the +root –sem condition, and 20 target words were tested under the +root +sem condition. Thirty-six target words were tested in each of the other conditions. Targets and primes were all nominal categories except for items in the morphologically related conditions, which were verbal forms (Table 8.2).

Twelve target words, primed with real words or nonwords, were used as practice items at the beginning of the experiment to familiarize the participant with the task. A total of 36 real word targets were primed with real words, and a total of 36 nonword targets were primed with nonwords. In addition, each participant heard a set of 72 incongruent target words with 36 real words primed with nonwords, and 36 nonwords primed with real words. Incongruent pairs were used in this experiment as fillers so as to prevent the participant from developing a systematic way of guessing whether the word is a real word, or not, based on the prime lexicality (El Hamdi 2018; Ussishkin et al. 2015). Congruent pairs were organized in three lists, and counterbalanced by priming condition using a Latin square so that each target word is primed with a different word in each list – ultimately, making the participant hear each target word only once. For incongruent pairs, counterbalanced lists were unnecessary since target words were primed under one priming condition, and, hence, were heard only once.

The use of nonwords in experiment 1 was designed to check for repetition/identity priming condition assuming that the participant would produce a faster response to nonword targets (Ussishkin et al. 2015). This suggests that repetition/identity priming condition produces a priming effect independently of the word’s lexicality. To create nonword stimuli, we first generated a list of possible, but unattested, roots in Tashlhit, which do not violate any structure constraints of the language. These included: OCP (Obligatory Contour Principle) with labials,

Table 8.2 Experiment 1. Priming conditions with sample prime-target pairs

	Identity	–root +sem	Unrelated	+root +sem	+root –sem
Prime1	ʃʃʳitʰ “a long piece of cloth used to tie”	izikr “rope”	afras “radish herb”		ʃʳitʰ “require”
Target1	ʃʃʳitʰ	ʃʃʳitʰ	ʃʃʳitʰ		ʃʃʳitʰ
Prime2	taʃʳtʰabʰt “broom”	ifsski “besom”	amzday “cemetery”	ʃʳtʰbʰ “sweep”	
Target2	taʃʳtʰabʰt	taʃʳtʰabʰt	taʃʳtʰabʰt	taʃʳtʰabʰt	

coronals, velars, and gutturals (Lahrouchi 2010, 2018), labial dissimilation /*msallam*/: [*nsallam*] “greet” (Boukous 1987; Bensoukas 1994, 2001, 2014), anteriority assimilation of sibilants /*skʃm*/: [*ʃkʃm*] “enter, causative” (Lasri 1991; Elmedlaoui 1992; Bensoukas 2004), and voicing assimilation /*ssugz*/: [*zzugz*] “take downstairs” (Lasri 1991). Second, we associated these roots with existing word patterns, concatenatively and nonconcatenatively. Examples of nonwords are: $\sqrt{n/}$: *tanʃla*, $\sqrt{ʃb}$: *taʃba*, *aʃlab*. All the created nonwords were vetted by two native speakers of Tashlhit to remove the stimuli, which were judged to be similar to an offensive word; or to resemble a real word.

Procedure

A female Tashlhit speaker recorded all the stimuli used in this experiment. The stimuli were recorded using a head-mounted Audio-Technica ATR3350 Omnidirectional Condenser Lavalier Microphone and a *TASCAM DR-40* digital recorder. All the words were repeated three times and we selected the stimuli with the clearest articulation. We demarcated, labeled, and extracted all the selected stimuli as sound.wav files, using Praat text grids and a Praat script (To label and extract the sound.wav files, we used the praat script written by Scott Jackson and last modified on 12/28/08.). To avoid having particular tones and speeds for real words and nonwords, we asked the participant to simultaneously record, both real words, and nonwords, which were randomly listed. We later calculated the mean duration for both real words and nonwords, and they were 600 ms and 591 ms, respectively, suggesting that, both real words, and nonwords, were recorded the same way by the participant. We then created prime-target sound.wav files through combining the recorded individual stimuli with an interstimulus interval of 150 ms (To create the prime-target sound.wav files, we used the Praat script, which was designed by Scott Jackson and edited by Jonathan Geary. We retrieved the script from <https://github.com/dbqpdb/auditoryMaskedPrimingStimGenerator>).

In a quiet place, we used a laptop computer, running DMDX software (Forster and Forster 2003), to administer the experiment, which lasted 10–13 min. On each trial, participants heard a prime, a 150 ms interstimulus interval of silence, and then the target. We asked the participants to respond as quickly and as accurately as possible to whether the second word they heard is a real word in Tashlhit, or not, and to use a Logitech Gamepad F310 to enter their responses by pressing the button on the right for “YES,” and the button on the left for “NO.” All the participants wore Audio-Technica ATH-M40X headphones while running the experiment. A 3000 ms timeout was measured from target onset, and the participants who did not respond within 3000 ms after hearing the target were presented with a visual timeout feedback and then presented with the following prime-target pair. For feedback visualization, we designed the following system of responses: “ $\sqrt{}$ ” / “X” / “...” to refer to “correct,” “incorrect,” and “timeout.” This was done for the purpose of avoiding embarrassing the illiterate participants, who might not be able to read the traditional visual feedback. The participant was first asked to start with 12 practice items before they ran the experimental stimuli. The practice stimuli consisted of both real words and nonwords, and they were paired under the three conditions.

When the participants finished running the experiment, we orally asked him/her to answer an oral questionnaire to control the sociolinguistic factors that might have an impact on the results. We manually recorded the answers of the participants. The questions (Table 8.3) were used as factors in the models we used in our analysis (El Hamdi 2018).

Answers to this questionnaire were used to select the data for analysis. Data from participants with hearing problems were excluded because the participant might have had a problem hearing the word well. We also excluded data from participants who did not speak Tashlhit as their mother tongue, and those who did not regularly use Tashlhit at home. We wanted to make sure that the negative response given by the participant is related to the word itself and not to other factors such as the scarce use of Tashlhit by the participant.

Results

Twenty-four participants, out of 54, were removed for their lower accuracy rate than 75%. Incorrect responses were also removed from the data analysis. We report here the results, using the reaction times (RTs) measured from target onset (RT onset). We also analyzed RTs from target offset and the results were identical to those we got from RT onset. We used a linear-mixed-effects regression model in our analysis using R: A Language and Environment for Statistical Computing (R Core Team 2016), and a lme4 package (Bates et al. 2015).

We used a random intercept model, and we created a new model, which includes random slopes to account for a within-subject variation “(1 + Cond|Subject),” and another one that also includes a within-item variation “(1 + Cond|Item).” A comparison of each of the two by slope models with the random intercept model suggested that the use of random slopes did not significantly improve the fit of the model. Random slopes by-subject report an insignificant result ($\chi^2(14) = 1$, $p > 0.05$), and the model including random slopes by-item failed to converge. Therefore, we used the random intercepts model as the simplest one. The factors, or fixed effects, used in our model are age, gender, school, and priming condition, along with the random effects: Subject and Item.

Table 8.3 Questionnaire

Age:	How old are you?
Gender:	Male or female?
Mother tongue:	Which language is your mother tongue: Tashlhit/Arabic/Both?
Hearing:	Have you had any hearing problems?
Headphones:	Did you wear headphones during the experiment?
Home language:	Which language do you mostly use at home: Tashlhit/Arabic/Both?
Hand:	Are you left-handed or right-handed?
School:	Have you attended school? If yes, where did you go to?

We checked whether the random intercepts model met the linear mixed effects assumptions. The distribution of the residuals was close to normal. A Shapiro-Wilk test of normality confirmed the results ($W = 0.90008$, $p\text{-value} < 0.001$). Only the effects of Age and school were correlated but this does not affect the results obtained. Our interest was the priming conditions, and none of them were correlated, satisfying the assumption of collinearity. We also checked whether our model met the assumption of homoscedasticity and found that the residuals showed, more or less, a similar degree of deviation from the predicted values. Hence, our model was proven to meet all the assumptions of linear mixed effects.

Priming occurred in the identity condition ($\chi^2(1) = 45.94$, $p < 0.001$), a condition in which the prime and target were morphologically, but not semantically, related ($\chi^2(1) = 10.76$, $p < 0.01$), a condition where both morphology and semantics are shared ($\chi^2(1) = 28.42$, $p < 0.001$), and in the semantics condition ($\chi^2(1) = 13.19$, $p < 0.001$). There is also a significant priming effect for the identity condition with nonword trials ($\chi^2(1) = 26.53$, $p < 0.001$). The means, standard deviations, and error rates of RTs of real words by priming condition are reported in Table 8.4, and graphically represented in Fig. 8.2.

Table 8.4 Experiment 1. Real words: mean, standard deviation, and error rates of reaction times (RTs) by priming condition

Condition	Mean (ms)	Standard deviation	Error (%)
Identity	741.86	300.41	18.84
+root+sem	764	285.53	24.04
+root−sem	834.83	305.56	28.49
−root+sem	844.21	330.44	20.65
Unrelated	965.46	296.45	19.46

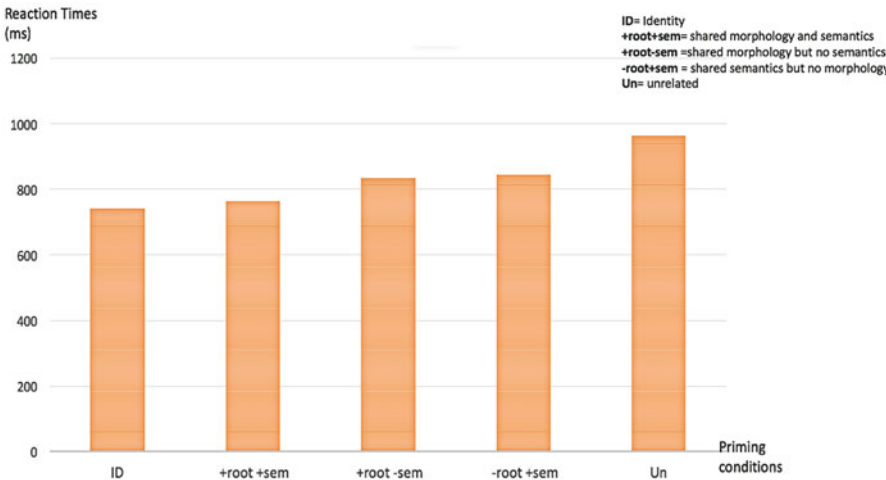


Fig. 8.2 Experiment 1. Mean reaction times of real words by priming conditions

Discussion

The results of this experiment point to a facilitatory effect when the words share semantic features. To account for the strictly morphological relation, one should control for other linguistic factors that may interfere with morphological word processing. Hence, to meet this requirement, we used pairs of real words that shared the root with an opaque semantic relation. Consequently, the results showed that a facilitatory priming effect occurred when the words shared only a transparent morphological relation (the root) with no semantic relatedness, suggesting that morphology alone can produce priming.

This is consistent with the results obtained in Semitic languages. In Arabic, root priming effect was robust, despite the fact that the semantic relation holding between the words is opaque (Boudelaa and Marslen-Wilson 2001, 2005, 2013, 2015; Schluter 2013; Gwilliams and Marantz 2015; Al Kaabi 2015). In the same vein, Hebrew studies also showed that words sharing only the primed each other (Deutsch et al. 1997, 1998, 2000b).

In addition, Tashlhit words sharing semantic features, but no root, also primed each other. These results confirmed the ones obtained in El Hamdi (2018), which reported a non-masked experiment, testing whether words that are semantically related primed each other. In order to determine whether there is a significant difference between the semantic and the morphological priming effect, we ran a Likelihood Ratio Test to compare the effect of words that are morphologically related, but semantically opaque (*ahddun-hiddn* “burnous-calm”), and words that are morphologically related, and semantically transparent (*taʃʔttʔabt-fʔttʔb* “broom-sweep”). We predicted that if semantic overlap had a significant effect, the morphological effect would have been significantly reduced if there was no semantic transparency between the words. The results showed that the effect of words sharing the root and a transparent semantic relation, and that of words sharing the root with no semantic feature overlap, are not significantly different: ($\chi^2(1) = 2.92, p > 0.05$). Hence, the two priming effects are indistinguishable, and the morphological effect is robust regardless of semantic transparency, suggesting that, both morphological structures and semantic features are lexically represented in the Tashlhit lexicon.

Similar results were attested in Semitic languages. In Hebrew, native speakers recognized the root as a morphological unit independent of any semantic overlap (Deutsch et al. 1997, 1998, 2000a). This is in keeping with the results obtained using incremental masked priming and lexical decision in Arabic visual word recognition (Boudelaa and Marslen-Wilson 2005). The latter study investigated the functionality of morphemes over distinct time points, and the results showed that root priming was obtained regardless of orthographic and or semantic relatedness. Unlike semantic and orthographic effects, root effects occurred at the four investigated SOAs (Stimulus-Onset Asynchrony). Semantic effect, on the other hand, only occurred at late SOA (80 ms). Moreover, the results showed that root priming had an equally strong effect with the words that have transparent semantic relation, and those words having an opaque one.

The significant effect of the semantic overlap in Tashlhit is not observed in Hebrew studies, in which semantic overlap is not as efficient as the morphological overlap. Deutsch et al. (1997) observed no priming when the words shared meaning, but no root, and that words sharing the root primed each other. The study tested semantic associates, showing that semantic relatedness did not facilitate lexical access. Although the authors did not deny that semantic priming might be noticed in different conditions, they contended that lexical access and target naming can only be facilitated when a morphological relation exists between the prime and target. Hence, root priming obtained, irrespective of semantic overlap in Hebrew, suggested that roots in Hebrew are lexical units independent of meaning. The results were confirmed using two different tasks: lexical decision and naming tasks. One account for this is that in Hebrew morphological decomposition theory is mandatory for the speaker, or reader, to process word recognition. If a root derivation is presented, the root is extracted as a default process by the native speaker, which allows for a quicker recognition of any other derivation of the same root, regardless of any semantic relatedness.

Experiment 2. Morphological and Phonological Priming Effect in Tashlhit

The results of the previous experiment showed a robust priming effect of the root, arguing for the important role morphology plays in word recognition. However, morphologically related words might also be phonologically related. For instance, the pairs *car-cars* and *friend-friendly* share both morphology and phonology; and pairs like *car-card* and *tin-tinsel* share phonology, but not morphology. Hence, different studies tend to investigate phonological and semantic factors independently of morphological factors.

In Hebrew studies, evidence has been presented for the facilitatory effect of the root independently of letters overlap. Letters that were not root letters, neither facilitate lexical access, nor target naming (Deutsch et al. 1997). This has been confirmed by Frost et al. 2005, who investigated to what extent form and/or orthography can facilitate lexical access. The results of their study showed that a facilitatory priming effect was observed when the prime and target shared the root, but not necessarily a sequence of letter; and that visual word processing is constrained by morphology, and not by phonology.

Arabic priming studies also showed that morphology plays an important role in lexical access, with emphasis on the nonsignificant role played by phonology. Words sharing two to three consonants with no morphological and semantic overlap (e.g., *samiidun* “semolina,” *famsun* “sun”) did not prime each other, suggesting that phonological overlap does not facilitate lexical access in Arabic (Boudelaa and Marslen-Wilson 2001, 2004a, 2013, 2015).

Using different modalities in English, the results did not support also the facilitatory effect of phonological structure. Using the auditory modality, the reaction times of words with more “neighbors” were long; that is, the RTs of words that share

phonological overlap were longer than words with fewer neighbors (Cluff and Luce 1990). On the other hand, Yates, Locker, and Simpson (2004) found evidence for phonological neighborhood in English using visual tasks. As a replication of the latter study, Grainger, Muneaux, Farioli, and Ziegler (2005) found different results, showing that the more orthographic features are shared between the words, the more chance phonological neighborhood density has in facilitating lexical access. In the same study, the authors also used a word identification task, and the results showed that phonology inhibits lexical access regardless of the number of orthographic neighbors. We contend that the results of the priming effects of the lexical decision task, shown in Grainger et al. (2005), are due to orthography and not to phonology proper and, hence, phonology, as tested in these studies, has an inhibitory effect in word recognition.

Similarly in the French language, Radeau, Morais, and Dewier (1989) examined phonological priming effect using lexical decision, and shadowing tasks in spoken word recognition. The results showed that no priming effects were observed in phonological cases sharing the first syllable (*/payyy/*, i.e., *PARURE* “ornament” and */pale/*, i.e., *PALAIS* “castle”), in those sharing the first phoneme (*/pale/*, i.e., *PALAIS* and */pule/*, i.e., *POULET* “chicken”), and in those changing the first phoneme (*/yule/*, i.e., *ROULER* “roll” and */pule/*, i.e., *POULET*). Phonological priming was also tested at different compression rates (35%, 40%, 50%, and 70%) in the same language, and the results were similar, showing no phonological priming effect (Kouider and Dupoux 2005). The results obtained in these studies suggest that phonological activation does not affect word processing, neither in Semitic, nor Indo-European, languages.

However, the absence of phonological priming is not confirmed by all priming studies. Phonological priming remains a debatable issue. The different interpretations of the effect of phonological priming might be due to the different priming techniques used in the experiments. This is illustrated in Moroccan Arabic where both phonology and morphology effects were found using supraliminal priming, but no phonology effects were obtained using subliminal priming. In the supraliminal priming experiment, morphology effects were “indistinguishable” from identity effects, suggesting that morphology is more robust than phonology, and that phonology activation facilitates lexical access in the later stages of word recognition in Moroccan Arabic (Schluter 2013). On the other hand, Halderman, Ashby, and Perfetti (2012) presented a review with a set of evidence that argues for the important role phonology plays in the early stages of word recognition process. Arguments have been presented from experimental studies on visual word recognition using different tasks, i.e., Eye movement and MEG studies.

In experiment 2 we tried to confirm whether the morphological effect obtained in experiment 1 is due to the shared morpheme, or just to the cluster of phonemes shared between the prime and target words. We attempted to test whether phonological factors yield a facilitatory priming effect, and we further examined whether the facilitatory effect of the root, obtained in experiment 1, is purely morphological, or phonological.

Method

This section provides details about the methodology adopted in experiment 2. It describes the subjects who participated in this experiment. It also depicts the experimental items and the priming conditions used in the experiment to test both the morphological and the phonological effect.

Participants

Seventy-four subjects participated in experiment 2 (42 females and 32 males). The age of participants ranged from 18 to 45 with the mean age of 26.71, and the median age of 25. Data from 28 participants were removed because their accuracy rate was lower than 75%; only data points from 45 participants were considered in the analysis. Incorrect responses were also removed from the analysis.

Materials

Each participant was exposed to a total of 140 target words including real words and nonwords in four main priming conditions: identity, phonology, morphology, and the control condition (Table 8.5). Morphologically related words shared the root morpheme. Primes were nouns and targets were trilateral verbs. In the remaining conditions, primes and targets were trilateral verbs. Each participant was exposed to 32 real words primed with real words, 32 nonwords primed with nonwords, and 64 targets used as fillers, 32 of which were real words primed with nonwords and 32 were nonwords primed with real words (Appendix B).

With regard to the phonology condition, we wanted to examine the phonological effect at all levels in which we test initial and final phonological overlap. Hence, words were divided into three subconditions: the change between the prime and target words was in the first phoneme (Phon1); the change was in the second phoneme (Phon2); and the change was in the third phoneme (Phon3).

Procedure

The procedure in experiment 2 was identical to that in experiment 1.

Table 8.5 Experiment 2. Priming conditions with sample prime-target pairs

Condition		Prime	Target
Identity		fsr	fsr “scatter, Verb”
Morphology		afsar “scatter, Noun”	fsr
Phonology	Phon1	nsr “blow one’s nose”	fsr
	Phon2	mt ^r “burry”	mgr “harvest”
	Phon3	y ^w r ^s “slaughter”	y ^w r ^d “lie down”
Unrelated/control		zfl “get nervous”	fsr

Results

Reaction Times (RTs) were measured from both target onset and target offset, and the results report no difference between the two. The reported results in what follows will be based on measuring response times from target onset.

We constructed the models in the same fashion as in experiment 1. The use of random slopes did not significantly improve the fit of the model. Random slopes by-subject reported a non-significant result ($\chi^2(20) = 8.64, p = 0.98$), and the model including random slopes by-item failed to converge.

Before we tested whether a priming effect occurred in any of the priming conditions, we examined whether the model met the linear mixed effects assumptions. A normal distribution of the residuals was observed and attested by the results of a Shapiro-Wilk test of normality ($W = 0.9439, p\text{-value} < 0.001$). No fixed effects were correlated/collinear. In addition, the data showed to be approximately equal across the range of the predicted values satisfying the assumption of homoscedasticity.

The means, standard deviations, and error rates of RTs by priming condition and by lexicality are reported in Table 8.6. A significant effect of priming condition was obtained with real words ($\chi^2(5) = 31, p < 0.001$). The experimental condition was significant for the identity ($\chi^2(1) = 28.10, p < 0.001$) and for the morphology ($\chi^2(1) = 8.90, p = 0.002$). No significant effect was observed for the three cases of the phonology condition with $\chi^2(1) = 1.07, p = 0.29$ in Phon1, $\chi^2(1) = 3.14, p = 0.07$ in Phon2, and $\chi^2(1) = 0.25, p = 0.61$ in Phon3. There was also a significant priming effect for the identity of nonwords ($\chi^2(1) = 7.13, p = 0.007$), as a result of the fast reaction to the “no” response when hearing the nonword target (Ussishkin et al. 2015).

Table 8.6 Experiment 3. Mean, standard deviation, and error rates of RTs by lexicality and priming condition

Condition	Mean (ms)	Standard deviation	Error (%)
Real words			
Identity	806.66	382.88	21.33
Root	870.51	444.87	24.41
Phon1	916.11	404.58	41.5
Phon2	885.34	408.43	37.44
Phon3	936.81	374.80	37.86
Unrelated	945.55	423.20	23.43
Nonwords			
Identity	872.46	385.35	22.28
Root	953.75	414.31	25.12
Phon1	953.4	408.91	45.43
Phon2	923.38	409.29	41.34
Phon3	974.25	450.11	50.32
Unrelated	954.95	404.62	24.40

Discussion

This experiment shows a facilitatory effect of the root, confirming the previous results of experiment 1. One possible interpretation of this result is that the effect obtained is due to the phoneme clusters. However, no phonological effect was obtained in this experiment even when the words shared the two first phonemes (Phon3). This suggests that what facilitates lexical access is not the shared phonological component, but rather the morphemic unit itself. One more argument is that the shared root is not necessarily presented in the two words as one cluster. For example, the verb *ak^wl* “stamp” primes the noun *akal* “land,” although a prefinal vowel separates the cluster of phonemes, but it does not prime the verb *ak^wr* “steal,” even when the two first phonemes are shared.

In the present study, we tested the degree to which phonological activation might have an impact on facilitating lexical access. We examined the phonological effect at three levels with trilateral verbs: a change in the first phoneme, a change in the second phoneme, and a change in the third phoneme. We predicted faster RTs yielding a phonological priming effect, at least when the change is in the third phoneme. However, no priming was obtained in either case, suggesting that the activation of the phonological structure does not facilitate lexical access in Tashlhit unless there is a transparent morphological relation between the prime and the target.

Previous studies on different languages showed different results with respect to initial and final phonological overlap. Cited in Radeau, Morais, and Segui (1995), Corina (1992) examined the role of syllable overlap in disyllabic words by using lexical decision task. The results demonstrated that priming occurred with initial syllable overlap, and with final syllable overlap, in disyllabic words, and with Rime overlap in monosyllabic words. This suggests that phonological activation is insensitive to syllable position. In Radeau et al. (1995), using auditory lexical decision and shadowing task with long and short interstimulus intervals in French, priming effect was obtained with words having final phonological overlap (sharing the last two phonemes), but not with words having initial phonological overlap (sharing the initial two phonemes), independently of word frequency. The authors concluded that priming, when prime-target pairs share final phonological overlap, does not tap the lexical level, but rather early stages of word processing (the prelexical level). The authors added that when the initial overlap is about more than one phoneme, there is a greater chance of obtaining an inhibitory effect when using shadowing task.

The different results obtained with respect to the phonological effect might also be due to the difference in the phonological shared properties. Phonological priming has also been investigated in terms of syllable role in comparison with shared phonemes. In English, phonological priming is obtained when the words shared the final syllable and not the final phoneme. For instance, *swimming* and *farming* share the final syllable *ming*, but not the morpheme. They primed each other due to their shared final syllable. No priming, on the other hand, was obtained with words like *breaking* and *smiling*, having the final syllables *king* and *ling*, respectively (Emmorey 1989). In Spanish, Carreiras and Grainger (2004) showed that the initial CV syllable yielded a fast response to the target word in Spanish *ju.nas-ju.nio*, but

not when the two words only shared the first three letters, and not the first syllable, *jun.to-ju.nio*, arguing for the important role of syllables in visual word recognition. This has been further examined in French in Carreiras, Ferrand, Grainger, and Perea (2005), where the role of both, the initial and final syllables, was examined in addition to the role of the first shared phoneme. Results showed that priming occurred only when there was an overlap in the initial syllable but not in the second one. The results were confirmed using two different tasks: lexical decision and naming suggesting that the processing of bi-syllabic words is sequential. The results are in support of the claim that phonological activation is early in reading (Lukatela and Turvey 1994). The priming effect of the first phoneme was only obtained in lexical decision but not in naming.

One might refer to the results obtained in this study as phonetic rather than phonological (Goldinger et al. 1992). However, knowing that in Tashlhit we might have syllabic consonants, and assuming the syllable structure in Jebbour (1996), the experimental stimuli used in our experiment share different syllable structures. We have pairs of monosyllabic-monosyllabic (*frs-srs*), disyllabic-monosyllabic (*kfm*, *k^wrm*), disyllabic-disyllabic (*xmr-g^wmr*), and monosyllabic-disyllabic (*srm-sri*). Not only was the overlap of initial and final phonemes investigated, but also initial and final syllable overlap was examined. Yet, the results showed no phonological priming effect. These results provided further evidence, from new data, for the absence of phonological priming. We conclude that the morphological priming effect is insensitive to phonological overlap, whether it is initial, or final, contrary to the studies that claim the dependent morphological effect on the phonological overlap (Gonnerman et al. 2007).

General Discussion

The results of these two experiments confirm the hypothesis that the root facilitates lexical access in Tashlhit, supporting the decomposition theory in which the root is extracted by the native speaker and allows a fast recognition of any derivation of the root in question. These results are consistent with the model of the lexicon in which roots are the lexical units that link all derivations, sharing the same root morpheme. In this study, we tested whether the facilitatory priming effect obtained in these experiments is a strictly morphological effect by controlling for semantic factors in experiment 1, and for phonological factors in experiment 2. The results of both experiments confirmed that it is the root that facilitates lexical access regardless of any other linguistic factor.

The strong effect of the root we obtained in experiment 1 is consistent with the results obtained in a number of priming studies in Arabic. No difference was found between words sharing a root with a transparent semantic relation, and those sharing the root with an opaque semantic relation (Boudelaa and Marslen-Wilson 2001, 2005, 2013). The reaction time to words sharing the root, with an opaque semantic relation (*mataaṣun* “commodity”-*mutṣatun* “pleasure”), was as fast as the one to words sharing the root with a transparent semantic relation (*mumtiṣun*

“enjoyable”-*mutṣatun* “pleasure”), suggesting that the effect of morphology was strong regardless of the transparency of semantic relation. In another study, morphological priming was shown to be robust regardless of the opacity or transparency of the semantic relation, even when the prime and target belonged to different grammatical categories (*ṣaqlun* “mind” [Noun]-*taṣaqqala* “be mindful” [Verb]) (Boudelaa and Marslen-Wilson 2015). Further evidence for the crucial role of the root in Arabic emerges from the auditory modality in which the results showed that the root constitutes an important unit that facilitates word recognition (Schluter 2013; Gwilliams and Marantz 2015).

One of the explanations for the lack of semantic priming effect provided by Schluter (2013) is that subjects might not have been given enough time to process Moroccan Arabic verbs at the semantic level – calling for future research to test for later semantic effects. On the other hand, we contend that the lack of semantic effect might be due to the type of semantic relation between the experimental stimuli. Following the same prime compression technique used by Schluter (2013), Al Kaabi (2015) investigated the role of roots in Emirati Arabic and found that roots facilitated lexical access in spoken word recognition; whereas, no semantic effect was obtained. In the same study, the strong effect of roots in Arabic was also supported by the results of an experiment on Standard Arabic using MEG recordings.

The role of morphology and semantics has also been investigated in Indo-European languages. Different results were obtained in different languages. In German, morphologically related words that are semantically opaque (*verstehen-stehen* “understand-stand”), and those that are semantically transparent (*aufstehen-stehen* “stand up-stand”), showed indistinguishable priming effect, claiming that morphological relatedness produced strong facilitation effect regardless of semantic relatedness (Smolka et al. 2009). On the other hand, in English, pairs sharing semantics primed each other (Radeau 1983). Some studies showed that root priming was obtained in English only when there was a semantic similarity between the prime and target. In late stages of visual word recognition, sharing morphology but no semantics (*successor-success*) did not facilitate lexical access in English. For example, *govern-governor* showed a priming effect, whereas *apart-apartment* showed no priming effect (Marslen-Wilson et al. 1994). The authors concluded that morphologically related words that have opaque semantic relations are organized distinctly in the lexicon regardless of their morphological relation. This is explained by Deutsch et al. (1997) as a consequence of considering semantic properties as the main clue to assess the morphological relations by English native speakers. The words *need* and *needy*, for example, would be assessed as morphologically related because they share meaning.

Semitic languages, on the other hand, have a different morphological system. Studies on Hebrew showed that native speakers can detect the morphological relation between words by recognizing the root independently of the semantic overlap (Deutsch et al. 1997, 1998). The difference between the results on English and Hebrew might be due to the different priming techniques used, or to the morphological system. Morphological relations are not sufficient for any storage,

and/or retrieval of lexical items in English, unlike in Hebrew (Deutsch et al. 1997). Further argument for the role of the root in Hebrew has also been provided in Deutsch et al. (2000b), which used the parafoveal preview information as another way to assess automatic and early stages of word recognition as masked priming paradigm. The results proved that the root facilitated both naming and lexical decision in Hebrew, and regardless of the spatial location of root letters, words sharing the root still primed each other, as has been observed in experiment 2 with Tashlhit data.

Pointing in the same direction, Boudelaa and Marslen-Wilson (2005) contend that the difference between the effects of morphology, semantics, and orthography cross-linguistically is due to the way in which native speakers of different languages encode meaning linguistically. This obviously reflects on the way the lexicon of different languages is organized. An interesting example was provided in Boudelaa and Marslen-Wilson (2005) about causativity. The latter is conveyed, either lexically (*teach*: cause to learn), syntactically (*make someone happy*), or morphologically (*widen*: cause to become wide). In Moroccan Arabic, causativity is expressed morphologically through gemination (*frrʿh* “make someone happy”), and in Tashlhit it is expressed through prefixal morphology (*sfrrʿh* “make someone happy”), affirming that morphological role differs cross-linguistically. In addition, the difference between English and Hebrew is not just about the morphological system but also about the phonological system. English words and morphemes are fully specified units at the morphological and phonological level, whereas in Hebrew and other Semitic languages, neither the root at the morphological level, nor the word pattern at the phonological level, are pronounceable units, and one cannot surface without the other. In Tashlhit, on the other hand, a large set of root units can be pronounceable forms (e.g., \sqrt{nkr} : /nkr/ “wake up”; \sqrt{lgr} : /lgr/ “lock”). However, the results of the present study confirm the significant role of the morphological decomposition theory, as is shown by Semitic languages.

Unlike experiment 1 in which morphological priming did not significantly increase when the words shared semantic features, in Southern Tunisian Arabic and Modern Standard Arabic, in an auditory-auditory lexical decision task, the reaction time to words sharing the root and meaning is faster than that of words sharing the root exclusively (Boudelaa and Marslen-Wilson 2013). This is consistent with the results obtained with English data using cross-modal lexical decision tests. In the latter, in morphologically related pairs, the more semantic similarity increases, the more priming effect increases, suggesting that priming magnitude is dependent on semantic and phonological transparency (Gonnerman et al. 2007). Bentin and Feldman (1990) also showed that words sharing roots and meaning showed a more robust repetition effect than words sharing the root but no meaning. In another study on form and meaning, Pastizzo and Feldman (2009) cited in Kiear and Joanisse (2011), the results showed that priming was greater in items sharing both form and meaning (*boat-float*) than in items sharing just form (*coat-float*), or just meaning (*swim-float*). In recent studies on English, using cross-modal priming and ERP (Event-Related Potentials), the results found that in late interval time, morphological

effects were more robust when the words shared both morphology and semantics claiming that “morphological relatedness is graded rather than absolute” (Kielar and Joanisse 2011). This is in keeping with the convergence of codes theory, which claims that morphology emerges from overlapping semantics, phonology, and orthography. However, it is important to note that the shared morpheme in Kielar and Joanisse (2011) is not the root but the suffix. Previous research on suffix priming effect showed that words sharing the same suffix did not prime each other (Emmorey 1989). Hence, the type of morphological relation is crucial in word recognition process. The morphological relation tested in Kielar and Joanisse (2011) is different from the morphological relation we investigated in the present study. Unlike suffixes, roots are showed to play a significant role in lexical access facilitation.

In controlling for phonological factors in experiment 2, the results were consistent with the results of previous studies on phonological priming in Arabic. One of the studies that used cross-modal priming paradigm showed that words having a phonological overlap did not prime each other, whereas words sharing morphology primed each other even if the two words lack any semantic or allomorphic shared properties (Boudelaa and Marslen-Wilson 2004b). In an auditory-auditory lexical decision task on Southern Tunisian Arabic and Modern Standard Arabic, phonological priming has been tested using words sharing two or three consonants but not the root. For instance, the words *mæzruub-mazirʕah* “in a hurry-farm” share the consonants (*mzr*) but they have the roots \sqrt{zrb} and $\sqrt{zrʕ}$, respectively. The results showed that no phonological priming occurred in this condition (Boudelaa and Marslen-Wilson 2013). A further argument for the inhibitory effect of phonology emerges from a later study by the same authors on Standard Arabic using cross-modal priming experiments. The study tested phonological effect using pairs sharing two to three consonants with no morphological and semantic overlap (*samiidun* “semolina”-*famsun* “sun”). The results showed that phonological effect did not occur in Arabic, and hence, has no effect on the root priming effect obtained (Boudelaa and Marslen-Wilson 2015).

The absence of phonological priming has also been attested in Indo-European languages. Marslen-Wilson et al. (1994) investigated to what extent phonological factors facilitate lexical access, and whether phonological properties are stored in the mental lexicon of English. In an experiment, where only morphology and phonology were tested, the words that were only phonologically related – but not morphologically related – did not prime each other, unlike words which were morphologically related. Two words sharing morphological features primed each other regardless of whether they shared phonological properties or not. Words like *friendly* and *friend* (sharing both morphology and phonology) would prime each other, but not words like *tinsel* and *tin* (sharing phonology but no morphology). This suggests that phonological overlap does not yield any priming effect unless some semantics or morphology is shared between the prime and target pairs. On just phonological priming, using auditory lexical decision task on English, the prime-target pairs were examined at the phonological condition at three levels: either they shared the first,

second, and third phonemes, the first and second phonemes, or just the first phoneme. The experimental stimuli were divided into high frequent and low frequent items, and no effect was noticed in the case of initial phonological overlap (Slowiaczek and Pisoni 1986).

Conclusion

The cross-linguistic difference in the organization of the lexicon might be due to the richness of the language morphology, or to the “form-meaning regularities” particular to the language. A morphologically rich language like German might have biased the representations of the structures in the lexicon. Furthermore, a language in which the overlapping of form and meaning are not idiosyncratic, but rather regular mappings of morphology and semantics, might have an influence on the lexical representations. We assume that Tashlhit has more cases of overlapping of form and meaning. In experiment 1, we had 20 words that shared both, the root and semantic features, and 16 words sharing just the root with no semantic overlap.

In the present chapter, we argued for morphological decomposition. Roots showed to be facilitating lexical access in Tashlhit, as has been shown in Hebrew, Arabic, and other languages. Hence, we argue for a model of lexicon organization in which root morphemes are extracted. The organization of roots in the model we advocate here is insensitive but related to semantic relatedness, and independent of phonology. The results of this study are consistent with the localist view of the mental lexicon, which assumes that morphologically related words are interconnected in the lexicon through the morpheme they share – as opposed to the nonlocalist view of the mental lexicon, which contends that morphological effects are due to the correlation between phonological, orthographic, and semantic effects, and not to morphology per se (Gonnerman et al. 2007; Kielar and Joanisse 2011). A theory on word processing would require a view on how native speakers of all, and different languages, access their lexicon. Hence, this study is a step forward for a wide scope of future research on a newly explored language from a psycholinguistic perspective.

Acknowledgments This work has been undertaken within the Partenariat Hubert-Curien Toubkal “Actions Intégrées” Volubilis project MA/14/311-Campus France N° 30285ZM entitled “De la nature et du rôle de la racine en amazighe: investigations sur la représentation mentale des mots”; and was supported by the Fulbright joint-supervision grant by MACECE (Moroccan-American Commission for Educational and Cultural Exchange).

Appendix A (Stimuli Used in Experiment 1)

See Tables 8.7 and 8.8

Table 8.7 Real words with morphologically related words having a transparent semantic relation

Target	–Root+Sem	+Root+Sem	Control
asqqal	azlag	sqq	ad ^{ad}
tajt ^t abt	ifski	jt ^t b	amzday
azlif	awri	zlf	asttix
ak ^w rram	ašmim	k ^w rm	tiftilt
afsaj	afukku	fsi	ammraz
awal	aqqur	sawl	taxmirt
igran	iyunan	grn	isdal
amkraz	awallu	krz	tazra
azil	anfa	zzil	ar ^z zz ^{um}
azuzwu	as ^{mmid}	zwu	annli
ay ^w rr ^{af}	lkas	y ^w r ^{af}	aknt ^{ur}
lhma	lhif	hmu	a ^{gmil}
zzif	a ^{zyay}	zif	agzzar
lbrih	la ^{jad}	brrh	anu
tamaya	taz ^{it}	may	ayar ^{as}
tazdm	takrrust	zdm	afud
irkan	abdduz	rku	tass ^{fs} aft
tamgra	tiwizi	mgr	taz ^{all} it
tag ^w mmimt	timqqit	ag ^w m	ffari ^z
ahwaf	nnd ^m	huf	ahlas

Table 8.8 Real words with morphologically related words having an opaque semantic relation

Target	–Root+Sem	+Root–Sem	Control
ifri	amda	fru	tiskrt
ah ^f mij	af ^r uy	h ^f jm	ibruri
ff ^r it ^t	izikr	ff ^r t ^t	afras
afullus	akazz ^{aj}	fls	imzizzl
ikru	tili	kru	tasa
asli	argaz	sli	atfl
az ^r u	ayulid	z ^r	addar
tirmt	allas	arm	iz ⁿⁱ
afiddun	aslhiam	fiddn	ahanu
tal ^{sint}	taybalut	l ^{sn}	amdlu
ff ⁿⁿ aq	algamu	ff ⁿⁿ q	az ^{alim}
tagant	lxla	gn	afus
ahfur	aq ^w zi	hfr	tayawsa
agru	alfsa	gru	usman
taffarnut	takatt	frn	Taqqajt

Appendix B (Real Words Used in Experiment 2)

Target	Phonology		Morphology	Control
frs	srs	Phon1	afras	rʃq
xmr	g ^w mr	Phon1	taxmirt	nyd
z ^b r ^c	s ^b r ^c	Phon1	az ^b bb ^c ar ^c	nqqʃ
ywd ^r	d ^r	Phon1	aydd ^r ar ^c	xtm
r ^d l	md ^r l	Phon1	art ^t al	zʃir
krz	mmrz	Phon1	amkraz	br ^c m
xdm	zdm	Phon1	axddam	r ^c z ^c m
zdr	bdr	Phon1	aʒdar	l ^b ʒ
fsr	nsr	Phon1	afsar	ʒʃil
ftl	ntl	Phon1	tiftult	rbh
mgr	mtr	Phon2	tamgra	ʒllx
sbʏ	stʏ	Phon2	asbbay	ʃrk
hfr	yfr	Phon1	ahfur	slʏ
rkm	rgm	Phon2	tirkmin	ʃt ^c n
nkr	ndr	Phon2	tankra	msl
gnu	gru	Phon2	tigni	nql
kʃm	k ^w rm	Phon2	akʃʃum	gzʒr
hgr ^c	hzz ^r	Phon2	ahggar ^c	tlf
hzzm	hʃʃim	Phon2	ahzzam	qrr ^b
lgr	lwr	Phon2	talgrawt	d ^c lm
srn	sri	Phon3	asram	ndm
frg	fm	Phon3	afrig	xl ^c d ^c
ak ^w l	ak ^w r	Phon3	akal	x ^w mʒ
arm	ars	Phon3	tirmt	fiddn
brrh	brrd	Phon3	lbrih	xnng
k ^w ms	k ^w mz	Phon3	ak ^w mmis	fd ^b h
y ^w r ^c s ^c	y ^w r ^c d ^c	Phon3	tiyr ^c s ^c i	nhs
syl	sʏi	Phon3	asʏal	k ^w rf
fr ^c ʃ	fr ^c ʒ	Phon3	lfr ^c aʃ	zdy
nʒh	nʒm	Phon3	nnaʒah	rkkz
zlf	zly	Phon3	azlaf	d ^c mn

References

- Al Kaabi M (2015) The anatomy of Arabic words: the role of the root in representations and processing. Doctoral dissertation, New York University
- Ameur M, Bouhjar A, Boukhris F (2004) Initiation à la langue amazigh. Institut Royal de la Culture Amazighe, Rabat
- Aronoff M (1994) Morphology by itself: stems and inflectional classes, vol 22. MIT Press, Cambridge, MA

- Balota DA, Pilotti M, Cortese MJ (2001) Subjective frequency estimates for 2,938 monosyllabic words. *Mem Cogn* 29(4):639–647
- Basset A (1952/2012) *La langue berbère*. Oxford University Press/International African Institute, London
- Bat-El O (1994) Stem modification and cluster transfer in Modern Hebrew. *Nat Lang Linguist Theory* 12(4):571–596
- Bates D, Kliegl R, Vasishth S, Baayen H (2015) Parsimonious mixed models. arXiv preprint arXiv:1506.04967
- Becker CA (1980) Semantic context effects in visual word recognition: an analysis of semantic strategies. *Mem Cogn* 8(6):493–512
- Benmamoun E (2003) The role of the imperfective template in Arabic morphology. *Lang Acquisit Lang Disord* 28:99–114
- Bensoukas K (1994) *Tashlhit Agentive Nouns- An Optimality-Theoretic Approach*. D.E.S. dissertation, FLHS, Rabat
- Bensoukas K (2001) Stem forms in the nontemplatic morphology of Berber. Doctoral dissertation, Mohammed V University, FLSH, Rabat
- Bensoukas K (2014) Featural dissimilation in Tashlhit: avoiding the repetition of [labial] and [round]. Mohammed V University, FLSH, Rabat
- Bensoukas K (2015a) Bu-nouns in Tashlhit: an oft-overlooked complex morphosyntactic corpus. *Corpus* 14:165–188
- Bensoukas K (2015b) Expressing ownership in Tashlhit: phrasal affix(ation) vs. bound word(hood). *Asinag* 10:11–38
- Bensoukas K (2016) Id-pluralization in Tashlhit: definitely not a root-and-pattern morphology! To appear in proceedings of *La Catégorisation Grammaticale en Amazighe*. IRCAM, Rabat
- Bensoukas K (2018) Against the consonantal root in Tashlhit. *Asinag* 13:53–90
- Bentin S, Feldman LB (1990) The contribution of morphological and semantic relatedness to repetition priming at short and long lags: evidence from Hebrew. *Q J Exp Psychol* 42(4):693–711
- Blumstein SE, Milberg W, Shrier R (1982) Semantic processing in aphasia: evidence from an auditory lexical decision task. *Brain Lang* 17(2):301–315
- Boudelaa S, Marslen-Wilson WD (2001) Morphological units in the Arabic mental lexicon. *Cognition* 81(1):65–92
- Boudelaa S, Marslen-Wilson WD (2004a) Abstract morphemes and lexical representation: the CV-skeleton in Arabic. *Cognition* 92(3):271–303
- Boudelaa S, Marslen-Wilson WD (2004b) Allomorphic variation in Arabic: implications for lexical processing and representation. *Brain Lang* 90:106–116
- Boudelaa S, Marslen-Wilson WD (2005) Discontinuous morphology in time: incremental masked priming in Arabic. *Lang Cogn Proc* 20:207–260
- Boudelaa S, Marslen-Wilson WD (2013) Morphological structure in the Arabic mental lexicon: parallels between standard and dialectal Arabic. *Lang Cogn Proc* 28(10):1453–1473
- Boudelaa S, Marslen-Wilson WD (2015) Structure, form, and meaning in the mental lexicon: evidence from Arabic. *Lang Cogn Neurosci* 30(8):955–992
- Boukous A (1987) *Phonotactique et domaines prosodiques en berbère (parler tachelhit d'Agadir, Maroc)*. Doctoral dissertation, Université Paris 8
- Boukous A (2009) *Phonologie de l'amazighe*. Institut royal de la culture amazighe, Rabat
- Boumalk A (2004) *Manuel de conjugaison du tachelhit (langue berbère du Maroc)*. Editions L'Harmattan, Paris
- Cantineau J (1950) Racines et schèmes. *Mélanges offerts à William Marçais*. G-P. Maisonneuve, Paris, pp 119–124
- Carreiras M, Grainger J (2004) Sublexical representations and the 'front end' of visual word recognition. *Lang Cogn Proc* 19(3):321–331
- Carreiras M, Ferrand L, Grainger J, Perea M (2005) Sequential effects of phonological priming in visual word recognition. *Psychol Sci* 16(8):585–589

- Chaker S (1990) La parenté chamito-sémitique du berbère: un faisceau d'indices convergents. *Etudes Docu Berbères* 7:28–57
- Chaker S (2008) Un standard berbère est-il possible? Entre réalités linguistiques et fictions sociolinguistiques. *Revue Etudes Berbères* 5:77–89
- Cluff MS, Luce PA (1990) Similarity neighborhoods of spoken two-syllable words: retroactive effects on multiple activation. *J Exp Psychol Hum Percept Perform* 16(3):551–563
- Coleman J (2000) The phonetics and phonology of Tashlhiyt Berber syllabic consonants. *Trans Philol Soc* 99(1):29–64
- Corina DP (1992) Syllable priming and lexical representations: Evidence from experiments and simulations. *Proceedings of the 14th annual conference of the cognitive science society*, pp 779–784
- Dell F, Elmedlaoui M (1985) Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *J Afr Lang Linguist* 7(2):105–130
- Dell F, Elmedlaoui M (1988) Syllabic consonants in Berber: some new evidence. *J Afr Lang Linguist* 10(1):1–17
- Dell F, Elmedlaoui M (1992) Quantitative transfer in the nonconcatenative morphology of Imdlawn Tashlhiyt Berber. *J Afroasiat Lang* 3:89–125
- Dell F, Elmedlaoui M (2013) Syllables and gemination in imperfective stems in Tashlhiyt Berber. *Brills J Afroasiat Lang Linguist* 5(1):1–34
- Deutsch A, Frost R, Forster KI (1997) What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *J Exp Psychol Learn Mem Cogn* 23(4):829–856
- Deutsch A, Frost R, Forster KI (1998) Verbs and nouns are organized and accessed differently in the mental lexicon: evidence from Hebrew. *J Exp Psychol Learn Mem Cogn* 24(5):1238–1255
- Deutsch A, Frost R, Forster KI (2000a) Decomposing morphologically complex words in a nonlinear morphology. *J Exp Psychol Learn Mem Cogn* 26(3):751–765
- Deutsch A, Frost R, Pollatsek A, Rayner K (2000b) Early morphological effects in word recognition in Hebrew: evidence from parafoveal preview benefit. *Lang Cogn Proc* 15(4–5):487–506
- Dupoux E, Mehler J (1990) Monitoring the lexicon with normal and compressed speech: Frequency effects and the prelexical code. *Journal of Memory and Language* 29(3):316–335
- El Hamdi F (2018) The presence of vowels in Tashlhit root structure: evidence from the free state and the construct state of the noun. *Asinag* 13:113–133
- El Moujahid E (1979) Présentation des phonèmes de la langue tamazighte: Le tachelhiyt d'Ighrem. *Traces* 2:52–78
- El Mountassir A (2003) Dictionnaire des verbes tachelhit-français:(parler berbère du sud du Maroc). Editions L'Harmattan, Paris
- Elmedlaoui M (1992) Aspects des représentations phonologiques dans certaines langues chamito-sémitiques. Doctoral dissertation, FLSH, Rabat
- Emmorey KD (1989) Auditory morphological priming in the lexicon. *Lang Cogn Proc* 4(2):73–92
- Forster KI, Davis C (1984) Repetition priming and frequency attenuation in lexical access. *J Exp Psychol Learn Mem Cogn* 10(4):680–698
- Forster KI, Forster JC (2003) DMDX: a windows display program with millisecond accuracy. *Behav Res Methods Inst Comput* 35(1):116–124
- Frauenfelder UH, Komisarjevsky Tyler L (1987) The process of spoken word recognition: an introduction. *Cognition* 25(1–2):1–20
- Frost R, Bentin S (1992) Reading consonants and guessing vowels: Visual word recognition in Hebrew orthography. *Advances in psychology* 94: 27–44
- Frost R, Kugler T, Deutsch A, Forster KI (2005) Orthographic structure versus morphological structure: principles of lexical organization in a given language. *J Exp Psychol Learn Mem Cogn* 31(6):1293–1326
- Galand L (1984) Le comportement des schèmes et des racines dans l'évolution de la langue: exemples Touaregs. In: Bynon J (ed) *Current progress in Afroasiatic linguistics*. John Benjamins, Amsterdam, pp 305–315

- Goldinger SD, Luce PA, Pisoni DB, Marcario JK (1992) Form-based priming in spoken word recognition: the roles of competition and bias. *J Exp Psychol Learn Mem Cogn* 18(6): 1211–1238
- Gonnerman LM, Seidenberg MS, Andersen ES (2007) Graded semantic and phonological similarity effects in priming: evidence for a distributed connectionist approach to morphology. *J Exp Psychol Gen* 136(2):323–345
- Grainger J, Muneaux M, Farioli F, Ziegler JC (2005) Effects of phonological and orthographic neighbourhood density interact in visual word recognition. *Q J Exp Psychol* 58(6):981–998
- Guerssel M, Lowenstamm J (1996) Ablaut in classical Arabic measure I active verbal forms. In: Lecarme J, Lowenstamm J, Shlonsky U (eds) *Studies in Afroasiatic grammar*. HAG, The Hague, pp 123–134
- Gwilliams L, Marantz A (2015) Non-linear processing of a linear speech stream: The influence of morphological structure on the recognition of spoken Arabic words. *Brain and language* 147: 1–13
- Gwilliams LE, Monahan PJ, Samuel AG (2015) Sensitivity to morphological composition in spoken word recognition: evidence from grammatical and lexical identification tasks. *J Exp Psychol Learn Mem Cogn* 41(6):1663–1674
- Halderman LK, Ashby J, Perfetti CA (2012) An early and integral role in identifying words. In: *Visual word recognition volume 1: models and methods, orthography and phonology*, vol 1. Psychology Press, New York, pp 207–228
- Hammond M (1988) Templatic transfer in Arabic broken plurals. *Nat Lang Linguist Theory* 6(2): 247–270
- Iazzi EM (1991) *Morphologie du verbe en tamazight (Parler des Aït Attabn Haut-Atlas Central): Approche prosodique*. DES dissertation, Mohammed V University-FLSH, Rabat
- Idrissi A (2001) *Toward a root-and-template approach to shape-invariant morphology*. Doctoral dissertation, Québec University, Montréal
- Jebbour A (1988) *Processus de formation du Pluriel nominal en Tamazight. (Tachelhit de Tiznit)- Approche non-Concaténative*, D.E.S. dissertation, FLSH, Rabat
- Jebbour A (1996) *Morphologie et contraintes prosodiques en berbère (tachelhit de Tiznit): analyse linguistique et traitement automatique*. Doctoral Dissertation, Mohammed V University, FLSH, Rabat
- Kielar A, Joanisse MF (2011) The role of semantic and phonological factors in word recognition: an ERP cross-modal priming study of derivational morphology. *Neuropsychologia* 49(2):161–177
- Kouider S, Dupoux E (2005) Subliminal speech priming. *Psychol Sci* 16(8):617–625
- Kunde W (2004) Response priming by supraliminal and subliminal action effects. *Psychological Research* 68: 91–96
- Laabdellaoui R, Boumalk A, Iazzi EM, Souifi H, Ansar K (2012) *Manuel de conjugaison de l'amazighe*. IRCAM, Rabat
- Lahrouchi M (2004) *Gémination, réduplication et gabarits dans un langage secret du berbère tachelhit*. JEL [Domaines], Nantes, pp 7–12. <halshs-00187602>
- Lahrouchi M (2010) On the internal structure of Tashlhiyt Berber triconsonantal roots. *Linguist Inquiry* 41(2):255–285
- Lahrouchi M (2018) La racine consonantique en amazighe: construit théorique ou réalité psychologique. *Asinag* 13:135–150
- Lasri A (1991) *Aspects de la phonologie non-linéaire du parler berbère: Chleuh de tidli*. Doctoral dissertation. Université de la Sorbonne Nouvelle, Paris III
- Longtin CM, Meunier F (2005) Morphological decomposition in early visual word processing. *J Mem Lang* 53(1):26–41
- Longtin CM, Segui J, Halle PA (2003) Morphological priming without morphological relationship. *Lang Cogn Proc* 18(3):313–334
- Lukatela G, Turvey MT (1994) Visual lexical access is initially phonological: 2. Evidence from phonological priming by homophones and pseudohomophones. *J Exp Psychol Gen* 123(4): 331–353

- Marslen-Wilson W, Tyler LK, Waksler R, Older L (1994) Morphology and meaning in the English mental lexicon. *Psychol Rev* 101(1):3–33
- McCarthy JJ (1981) A prosodic theory of nonconcatenative morphology. *Linguist Inquiry* 12(3):373–418
- McCarthy JJ, Prince AS (1990) Foot and word in prosodic morphology: the Arabic broken plural. *Nat Lang Linguist Theory* 8(2):209–283
- Milin P, Smolka E, Feldman LB (2017) Models of lexical access and morphological processing. *The handbook of psycholinguistics*, pp 240–268
- Prunet JF, Béland R, Idrissi A (2000) The mental representation of Semitic words. *Linguist Inquiry* 31(4):609–648
- R Core Team (2016) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Radeau M (1983) Semantic priming between spoken words in adults and children. *Can J Psychol* 37(4):547–556
- Radeau M, Morais J, Dewier A (1989) Phonological priming in spoken word recognition: Task effects. *Memory and Cognition* 17(5):525–535
- Radeau M, Morais J, Segui J (1995) Phonological priming between monosyllabic spoken words. *J Exp Psychol Hum Percept Perform* 21(6):1297–1311
- Rastle K, Davis MH, Marslen-Wilson WD, Tyler LK (2000) Morphological and semantic effects in visual word recognition: a time-course study. *Lang Cogn Proc* 15(4–5):507–537
- Schluter KT (2013) Hearing words without structure: subliminal speech priming and the organization of the Moroccan Arabic lexicon. Doctoral dissertation, The University of Arizona
- Ślowiacek LM, Pisoni DB (1986) Effects of phonological similarity on priming in auditory lexical decision. *Mem Cogn* 14(3):230–237
- Smolka E, Komlosi S, Rösler F (2009) When semantics means less than morphology: the processing of German prefixed verbs. *Lang Cogn Proc* 24(3):337–375
- Tobin Y (1990) A combinatory phonology of the Hebrew triconsonantal (CCC) root system. *La Linguistique* 26:99–114
- Ussishkin A (1999) The inadequacy of the consonantal root: modern Hebrew denominal verbs and output–output correspondence. *Phonology* 16(3):401–442
- Ussishkin A, Twist A (2009) Auditory and visual lexical decision in Maltese. In: *Introducing Maltese linguistics: selected papers from the 1st international conference on Maltese linguistics*. John Benjamins, Amsterdam, pp 233–249
- Ussishkin A, Dawson CR, Wedel A, Schluter K (2015) Auditory masked priming in Maltese spoken word recognition. *Lang Cogn Neurosci* 30(9):1096–1115
- Winther Balling L, Harald Baayen R. (2008) Morphological effects in auditory word recognition: Evidence from Danish. *Language and Cognitive processes* 23(7–8):1159–1190
- Yates M, Locker L, Simpson GB (2004) The influence of phonological neighborhood on visual word perception. *Psychon Bull Rev* 11(3):452–457



Shaping Amazigh Identity: The Case of the Netherlands

9

Jan Jaap de Ruiter

Contents

Introduction	195
Moroccans in the Netherlands	197
Leading Figures with an Amazigh Background in the Netherlands	199
Amazigh Activists	200
Amazigh Associations in the Netherlands	201
Research on Berber	203
The Language Issues	204
The Berber Community in the Netherlands: A Case of Transnationalism?	205
Outlook	206
References	208

Introduction

Moroccan emigration to the Netherlands started in the 1960s with individual workers reaching the country, mostly via France and Belgium, and continued with the recruitment of workers, which was based on a treaty between both kingdoms in 1969. Workers came, in particular, from the North of Morocco, which already had a history of work migration. Later, wives and children joined their husbands/fathers, and more children were born in the new country. Today (i.e., in 2021), there are 408,864 people in the Netherlands with a Moroccan background, consisting by now of three generations and a fourth one developing (Statline, bevolking kerncijfers). They form just over 2% of the population, which in November 2020 numbered 17,422,992 people (Statline, Bevolkingsontwikkeling; maand en jaar).

J. J. de Ruiter (✉)
Tilburg University, Tilburg, The Netherlands
e-mail: jj.deruiter@tilburguniversity.edu

The “Moroccan community” is very diverse, consisting of the different generations with their varying levels of command of Dutch and of the “original” languages of Morocco: the varieties of Arabic and Berber. Dutch “Moroccans” use their diverse languages on different levels, with the younger generations overall fluent in Dutch and not or hardly using the original mother tongues (the varieties of Arabic or Berber) anymore, while the older generation is still having trouble with Dutch, expressing themselves mostly in the mother tongues. But there is the danger of generalization. There are also quite some first-generation migrants who speak Dutch fluently while some young newcomers, Moroccan migrants entering the Netherlands in the tens and twenties of this century, hardly speak the language yet (Slootman 2018).

In any case, the original language background of “the Moroccans” is very diverse, be they speakers of Arabic dialects, known as *Darija*, and/or of the Berber or Amazigh varieties spoken in Morocco, mainly of the Northern Riffian variety, *Tarifit*, and to a lesser extent of the Tamazight varieties spoken in the center of the country, and of *Tashelhit*, the southern variety of Amazigh. It is estimated that around one third of Dutch Moroccans has an Arabic language background and the others an Amazigh one (de Ruiter 1989, 2019).

Another variable playing a role in the Moroccan community is religion. In general, “Moroccans” are more serious in abiding by the rules of their religion, Islam, than the average Christian Dutchman. There is also a faction within the community that has radicalized and joined extremist Islamic movements like (at the time) Islamic State (van San 2018). Most Dutch jihadis joining that organization were from a Moroccan background. Also, Moroccans are active in the Salafi movement in the Netherlands, which counts around 60,000 members, not all of them Moroccans. Some are Salafists with a Turkish background or Dutch Salafi converts (de Koning 2012–2013).

Then there are artists, writers, poets, and politicians with a Moroccan background who profile themselves in their specific fields of interest. These persons may or may not cherish religious values. Some generally abide with the rules of Islam; others are like Christians who only go to church for Christmas or, in the case of Islam, only celebrate the *Id al-Fitr* without fasting too much. And some of them openly declare that they are not Muslim anymore, like lauded writer Hafid Bouazza who passed away in 2021 (Caubet 2006).

This chapter focuses on the Amazigh community in the Netherlands, posing the question if, and if yes, to what extent, it succeeds in voicing its Amazighness, both in the national public debate, and within the Moroccan community itself. The subject is discussed against the background of what is currently taking place in Morocco, where Amazigh has recently been recognized and is mentioned in article 5 of the latest version of the Constitution (2011), which reads as follows:

Arabic remains the official language of the state. The State works for the protection and development of the Arabic language, as well as the promotion of its use. Likewise, Amazigh is an official language of the state, as a heritage common to all Moroccan, without exception.

Remarkably, Amazigh is not mentioned as an independent language but in tandem with Arabic. Nevertheless, the language is mentioned as an official

language of the Moroccan kingdom. But the position of Amazigh is not stabilized yet. One reason is the Moroccan Institute of Amazigh Culture (IRCAM), which in 2001 was to promote Berber language and culture in Morocco, and which is now facing a possible devaluation in the Moroccan language landscape due to a reform of the organizational structure of this kind of institute in the Kingdom, making IRCAM part of the new *Conseil national des langues et de la culture marocaine*. The fact that IRCAM will no longer be an independent institute may lead to a weakening of its position in Morocco's language policies, but time will tell. Another issue that the Moroccan community in the Netherlands faces is the force of populism, headed unmistakably by Geert Wilders, leader of the Dutch Party for Freedom, which is openly opposed to Muslims and (therefore) to Moroccans (de Ruiter 2012a, b). Then there is the "new kid on the block," Forum for Democracy leader Thierry Baudet, whose party pays lip service to inclusivism, but whose party program breathes the air of xenophobia and Islamophobia akin to Wilders'. Moroccans in the country are confronted daily with populist-inspired attacks, the only reason being that they are part of Dutch society.

This chapter is divided as follows. First, a general description of the societal position of people with a Moroccan background in the country is presented. This section is followed by a description of the position of the Amazigh layer of the Moroccan community in the Netherlands, focusing first on key figures in the community, and the activists among them. Thereafter, it will speak of their sense of organization within the Moroccan community. Next, the chapter discusses academic research done on Berbers in the Netherlands and then offers an overview of what this author calls the language issue: should Berber be taught in primary education in the Netherlands. Finally, an analysis will be presented to answer whether the Moroccan community, and thus the Amazigh community, should be viewed as a case of transnationalism.

Moroccans in the Netherlands

When the first laborers of Moroccan origin came to the Netherlands they were called "guest workers". The idea was that these workers would stay in the country temporarily and return home after gathering some capital to help build their lives "at home": the opposite happened though. Wives and children followed the laborers, and more children were born. This led to the Dutch government reconsidering its migration policies. With the *Minorities Act* of 1983, it was now considered a multicultural society, where the new migrants, most of them Moroccans and Turks, were expected to integrate into Dutch society (Entzinger 2006). Since then, Dutch policies have always focused on integration. But gradually a change in the concept emerged. In the eighties and nineties of the last century the government motto was "integration while preserving the original culture." Later, though, especially in the new century, integration was aiming for assimilation. Immigrants were

expected to assimilate themselves into Dutch values and norms. This led to a still-persistent crisis in Dutch society. A vigorous discourse developed as to the adaptability of Islam, the religion of most Moroccans and Turks, to Dutch values (Entzinger 2014). It was the then leader of the Dutch Liberal Party, Frits Bolkestein, who in his Lausanne lecture in 1991 raised this claim: he claimed that it would be far-fetched to think Islam and its values could adapt to Western ones. This discourse came up at the same time in which “guest workers” were being referred to as Muslims. The discourse initiated by Bolkestein was also adopted by the populist politician Pim Fortuyn, who was killed in 2002 by a leftist activist and whose party LPF (List Pim Fortuyn) a few days later had major success in the Dutch parliamentary elections, winning 26 seats out of 150. The torch of populism was then taken over by Geert Wilders, former Liberal Party MP, who at one point held no less than 24 seats in parliament, and now (2021) 17. In the 2017 elections Forum for Democracy leader Thierry Baudet entered Parliament with two seats, and in the 2021 elections he had eight. He promotes a political program of xenophobia and utter nationalism, vowing, among other things, to leave the European Union and the Euro currency. Together with the events of 9/11 and the killing in 2004 of the Dutch filmmaker Theo van Gogh by Mohammed Bouyeri, an Islamic extremist with a Moroccan background, the situation of Dutch citizens of Moroccan origin and others with an Islamic background has worsened.

Relations between Morocco and the Netherlands, which date back to 1610 (Bos and Fritschy 2006) and which were the first formal relations Morocco entered with a European nation, have been under stress soon after Moroccan workers first arrived in the Netherlands. One of the first issues was which language should be taught to Moroccan children in the Dutch primary schools in their separate lessons “Education in Original Language and Culture (EOLC).” Furthermore, the Netherlands believes that it can criticize political issues that Morocco regards as its own internal affairs. The Netherlands has always criticized Morocco’s human rights policies, both under King Hassan II and under the current King Mohammed VI. The Netherlands was also critical on the strong bonds the Moroccan government wanted and still wants to maintain with its citizens living abroad. Also, the issue of dual nationality stood as an obstacle between both countries, as the Dutch government does not recognize people having two nationalities but it cannot but accept this reality. The Moroccan government refuses to allow its citizens to give up their nationality. The Dutch government pressed Morocco regularly to abide by human rights. An example is the Hirak rebellion in the Rif region in 2016–2017, which was severely suppressed by the Moroccan authorities. Finally, there is the refusal of the Moroccan government to take back failed asylum seekers, which in turn led to the formation of a group of illegal Moroccans (and similarly with other nationalities like Tunisians and Egyptians) in the Netherlands. The stress between both countries found its apogee in the Covid-19 crisis in the spring of 2020 when Morocco refused to let Dutch Moroccans go back to the Netherlands, while allowing French Moroccans to return to France. Only after stiff negotiations these people were gradually allowed to return home.

Leading Figures with an Amazigh Background in the Netherlands

Halim el Madkouri (born in 1963) appeared regularly in the Dutch press as a specialist on Islamic extremism in the Netherlands, and on jihadis leaving for the then Caliphate in Syria. He is known for his critical opinions on the Moroccan Makhzen (central government), but more radical activists criticize him for his attenuating tone toward the same Moroccan government. The Amazigh website Arif News criticized el Madkouri for allowing an interview to the Moroccan news website *Hespress*, considered a “mouthpiece” of the Moroccan government. Arif News also alluded to charges against el Madkouri for financial malpractice when he was still working for the organization FORUM, an Institute for Multicultural Issues, which worked for the interests of migrants in the Netherlands. It was established in 1997 and stopped in 2015. In the end, El Madkouri was acquitted of the charges, and it seems meaningful that Arif News still wanted to mention the unpleasant affair, undoubtedly aiming to harm the activist who is a moderate, yet critical Amazigh voice in the Netherlands.

Ahmed Aboutaleb (born in 1961) started his career as a journalist for the local Radio Stad Amsterdam (Radio City of Amsterdam). In that capacity he interviewed me in September 1989 when I had defended my Ph.D. dissertation on the language situation of young Moroccans in the Netherlands (de Ruiter 1989). Later, he made it to the Labour Party as alderman for social affairs in the city of Amsterdam. Then for a short time (2007–2008) he held the office of State Secretary of Social Affairs in the fourth Balkenende government (2007–2010). He successfully applied for the mayor position of Rotterdam, a position that he has held since 2008. Aboutaleb can be regarded as a successful migrant’s son. His roots lie in a very poor part of the Rif region. Still, his position in the Netherlands is not undisputed as orthodox Muslims blame him for sharing a cause with the Dutch *kafir* government, while Riffian activists blame him for his contacts with the Moroccan government and the King. He made headlines when at the time of the Charlie Hebdo shooting in January 2015, he said that “if you cannot bear the irony of the Charlie Hebdo cartoons, you should just fuck off out of the country,” for which he was praised but also criticized, in particular by Syrian human rights organizations, which were unable to stop violent jihadis from Europe arriving in their country. Aboutaleb does not deny his Berber background, but it disappears a bit against the background of other more general political activities he must undertake as mayor of Rotterdam, and prominent member of the Labour Party. Ahmed Marcouch (born in 1969), also a Labour Party member and presently mayor of the city of Arnhem, like Aboutaleb from Moroccan origin and Berber, explores and expresses his Berber identity more than Aboutaleb does. He regularly had video messages recorded in Tarifit, in 2020, for example, on the dangers of the Covid 19 virus.

Abdelkader Benali (born in 1975) made his fame as a writer and public intellectual in the Netherlands. He is, by far, the most well-known writer-performer-artist with Moroccan-Amazigh roots in the country. He won several Dutch literature awards, among them the Libris Literature Award in 2003 and the E. du Perron

Award in 2009. Apart from an impressive list of novels, which were very well received and translated into foreign languages, Benali regularly tours the country in one man show: in 2017 with the show *Letter to my daughter*, and in 2020 *Caliph of the Netherlands*. He regularly participates in talk shows, in which he expresses his aversion against the ongoing hate rhetoric against Moroccans and Muslims. He also got irritated that he was labeled a migrant writer. He considers himself an artist without ethnic affiliations, although he cannot help giving voice to the despair Moroccans, Muslims, and non-Dutch citizens in general feel in a society where Muslim hate and xenophobia are gaining steam every day. Benali refers to his Berber background in his public performances but not exclusively, as he also tackles numerous other subjects of public debate. In 2005 the quality newspaper *NRC* published an opinion piece by him entitled “Why would Berbers stink” in which he refutes a number of prejudices on Berbers. The piece raised quite a debate in its aftermath. There was also a controversy in the Netherlands about the commemoration of 75 years of liberty after World War II in 2020. Jewish writer Arnon Grunberg, in a speech, compared the Jews then with the Moroccans now, and that sparked a national debate. Benali was invited to discuss the issue on the national talk show *Op1* on May 9, 2020. Joost Eerdmans, then leader of the local populist party *Leefbaar Rotterdam* (‘Livable Rotterdam’) spit out all possible prejudices about Moroccans on him. At a certain point the table was discussing the “question of the Moroccans,” ignoring the only Moroccan at the table, which led Benali to later state in a comment: “That is being a Moroccan in 2020 in the Netherlands: you are very visible but you can also, just like that, become invisible”. The irony of history was once again confirmed in a subsequent event. The National Committee May 4–5, which, each year, organizes the commemoration of the liberation of the Netherlands from Nazi occupation, invited Benali to give a lecture on the 76th national commemoration in 2021. After the news went public, reports surfaced in the media referring to allegedly anti-Semitic tweets from Benali – that he had previously sent out to the world. This led to a riot, and Benali’s withdrawal from giving the lecture. He didn’t want to be, he said, “a topic of conversation” at such an important national event.

Amazigh Activists

Mohammed Chacha (1955–2016) lived a considerable part of his life in the Netherlands. He can be labeled an activist of the Berber cause *par excellence*. He wrote in Arabic and Tarifit but less in Dutch, which explains why he isn’t that well known to the Dutch public. He was a stiff critic of the Moroccan regime and of Islam, which led to him temporarily being denied entry into Morocco, the country where he was buried in the end.

The Berber born cartoonist Abbttoy (born in 1963) regularly publishes cartoons depicting injustice and suppression in Morocco, particularly in Berberophone areas. His is a collection of cartoons called *Cartoons van een Berber* (2008) (*Cartoons of a*

Berber) in which he sketches his anger and frustration on the injustice taking place in Morocco, the suppression of the Berber language and culture, and also the rise of Salafi Islam, both in Morocco, and the Netherlands. In Abbtöy (2011), his second collection of cartoons appeared, called *De s van het kwaad* (*The s of evil*), in which Party for Freedom leader Geert Wilders is on the front cover. Asked about his presenting Mr. Wilders on the front page, knowing that he is the leader of an intolerant party that aims at chasing the Muslims out of the country, Abbtöy said: “Wilders is right and at the same time he is not. His crusade should be aimed at radicals instead of at ‘Islam’.”

Public intellectual and Berber Asis Aynan (born in 1980) is known as a columnist in several leading Dutch newspapers; he is also the author of several novels. The publication of the “Berber Library” was an initiative of Aynan’s as well. The Berber Library contained ten volumes of Berber writers translated into Dutch. For this series, I translated Libyan writer Ibrahim al Koni’s novel *Golddust* into Dutch (de Ruiter 2012c). Aynan is also known for having signed the petition “to get rid of the Moroccan nationality” in 2019, together with 11 other Dutchmen with Moroccan roots because “he wants to have nothing to do with the Moroccan government.” The petition met with wide public attention, but the goal hasn’t been reached so far. Of course, the action had a highly symbolic function. It is well known that Moroccans are, by law, unable to get rid of their home nationality, as Morocco wants to maintain the bonds with “les Marocains résidents à l’étranger,” for keeping a political eye on them and for making sure they continue sending money to their home country. In an interview with the Dutch daily *Trouw*, Aynan gave as the reason for his petition, “things are simple: I am not a Moroccan. I was born in the Netherlands. My parents were born during the Spanish protectorate and my grandparents were born during the Rif Republic [...] But the most important reason for me is that Morocco is a violent dictatorship.” In his novel *Een erwit maakt nog geen snert* (*One pea doesn’t make pea soup* 2020) Aynan discusses dual nationality of Dutch Moroccans, while, at the same time, explaining several misunderstandings regarding them.

Amazigh Associations in the Netherlands

Moroccan citizens have established numerous associations in the Netherlands. In their 2007 publication Van Amersfoort and Van Heelsum mention that there are 700 such associations. This section discusses those organizations, and websites that are relatively active and it focuses on those which were recently, in one way, or another, making the news (2019–2020). This is because there are also many associations (and their websites) that hardly show any activities or are not regularly kept up to date.

The Dutch language website arifnews.com states it is “an independent website whose goal is to provide its readers with mostly information and news about North Africa and ‘Arif’ (= the Rif area, JdR) in particular.” Foreman of this association is

“Amazigh Ayou” (not his real name). In an interview with the Dutch journal *De Kanttekening*, he stated that the Riffian activists may have their differences, but they all agree on the Moroccan state to be their common enemy: “The government has been eradicating these people (i.e., the Riffians JJdR), and their identities, for decades.” The website contains articles on the period of the Rif republic (1921–1926) and its legendary leader Abdelkarim al-Khattabi.

In July 2017, Socialist Party Parliamentarians Lilian Marijnissen, and Sadet Karabulut visited the northern Moroccan city of Al Hoceima, and the parental home of Nasser Zefzafi, the leader of the Riffian Hirak protest movement. Riffian activists like Amazigh Ayou appreciated this visit of solidarity after the uprisings in the Rif in 2016–2017, but the fact that both were also meeting with political leaders who were associated with the central Moroccan government, the so-called “Makhzen”, was criticized.

Hossniya el Yahyaoui, researcher at Leiden University, is chair of the Freedom and Human Rights Organization, whose aim is the promotion of human rights around the globe. The organization targets the situation in Morocco and other North African countries. El Yahyaoui is known as an activist for the cause of the Rif and is active in left-wing political circles.

In February 2020, seven Riffian organizations, the Bades Foundation, Stichting Izouran, Comité Mulay Mohand, Stichting Noemidia, Rif Alert, Vereniging Symphony, and Rifproject sent a letter to the members of the Dutch Parliamentary Committee on Foreign Trade and Development Cooperation who were to visit Morocco that same month. In this letter they asked the committee to raise the issue of human rights violations in the Rif. There is no information on the effect of this letter, as the visit never took place because of the Covid 19 pandemic. A case in point is the celebration of the Amazigh New Year 2969 (2019) in the city of Rotterdam, in which the Amazigh association Areef, from Al Hoceima, participated. The Dutch Amazigh website arifnews claimed that this trip was sponsored by the Moroccan government and it warned against the increasing influence of the Makhzen on Amazigh associations outside Morocco: this, in order to control them, and push them to more Morocco friendly attitudes. The fear of Morocco interfering in internal Amazigh affairs is general among Amazigh organisations.

The Dutch language website amazightimes.nl does not mention any names of owners or web masters. Under the link “about us” no information is provided and the columns it presents bear no names of authors. The links “abroad” and “North Africa” contain articles which are nearly all on Morocco. The site is very Moroccan government critical, if not hostile.

Then, there is the Dutch language website amazigh.nl, which is not well maintained, and the articles are anonymous and there are no names of editorial members. The site contains a page “stories and poems” in Riffian Amazigh written in the Roman alphabet. The website adrar.org provides material for learning Tarifit, as well as information on Berber lessons. The website is maintained by the cultural association Adrar and receives close to 3000 unique visitors per month. Another website devoted to Berber poetry is izranvantoen.nl.

Research on Berber

Leiden University has had a chair on Berber studies for many years. Harry Stroomer (1946) was the first to hold the position. His research focused on Tashelhit Berber, past and present. His works breathe, one might say, a traditional philological panache. He was succeeded by Maarten Kossmann (1966) in 2017, whose research, when it comes to Morocco, focuses on the Berber varieties in the north of the Kingdom, the Berber of Figuig included. Together, with Khalid Mourigh, mentioned below, he also paid attention to the slang of Moroccan youngsters in the Netherlands, focusing on Berber and Arabic borrowings in it. When accepting his post, he said in an interview: “Public discourse seems to be blind to the great diversity that characterizes this very community. One of the tasks of a professor of Berber to me seems to do justice to that diversity.” The academic output of the new professor of Berber shows, indeed, that he does pay attention to his diversity. He is quite positive about Morocco opening to Berber in recent years as well: “One sees information boards in Tifinagh everywhere in the country.” He notes that the mere mention of Berber in Morocco in the 1980s could lead to major problems with the authorities, an experience that I share, and that is no longer the case. One can criticize the current Moroccan authorities on many aspects in the country today but the fact is that the country is much more diverse, albeit on the surface only, than it was 40 years ago.

Khalid Mourigh (born in 1981) received his Ph.D. in linguistics at Leiden University in 2017 on a thesis that deals with the grammar of Ghomara Berber (Mourigh 2003). At present he publishes on street language of Moroccan youngsters, focusing among other issues on Amazigh and Arabic – appearing in that language. He translates Berber poetry and novels into Dutch. Mourigh can be classified as an academic, focusing on the cultural aspects of the Rif and Riffians in the Netherlands. He was also a member of the advisory board of the abovementioned “Berber Library.” In February 2021, Mourigh made his debut with a nonfiction novel about the migration of his grandfather (Mourigh 2021).

Abderrahman el Aissati (born in 1960), from a Berber background, paired his activist life to his academic pursuits. While affiliated with Tilburg University, he published on various aspects of Berbers in Morocco, and in the Netherlands. He also initiated Amazigh TV, which, finally, he had to end due to a lack of financial resources. The related YouTube channel has close to 13,000 subscribers. El Aissati is active in the internal Dutch Moroccan debate on issues related to Berber, both in festivals and conferences, in the Netherlands, Morocco, and elsewhere.

I have also paid attention to the role of Berber language and culture in numerous publications. My Ph.D. thesis (1989) presented the first overall image of the language situation of young Moroccans in the Netherlands. In it I gave an overview of the command and use by young Moroccans of their mother tongues, and Dutch. The general image was that the younger the informants were (there were four groups: 7, 11, 14, and 17 years old), the better was their command of Dutch, and the older they were, the better their command of Arabic and Berber. One must take into consideration that the research was executed in the 1980s, and the informants, therefore, members of the

one and a half generation (born in Morocco), or second generation (born in the Netherlands). The image that the research produced would not be valid anymore now, as in general young children from Moroccan origin, and born in the Netherlands, hardly have an active command of Arabic and Berber any more.

I executed sociolinguistic research in Morocco as well. In the 1990s and the first decades of the new century I had questionnaires distributed in Morocco among academic students. I wanted to establish the language profiles of these youngsters as I had done with young Moroccans in the Netherlands for the subject of my Ph.D. My research in Morocco led to the image that, in general, Standard Arabic enjoyed a high prestige, also among Berber students; yet, at the same time, the Berber identity was far from dead. As the language situation in Morocco itself is not the subject of this chapter, I will not dwell on my research done in the Moroccan Kingdom and refer interested readers to major publications on these themes (cf. De Ruiter 2001, 2006; Ruiter 2008; De Ruiter & Ziamari 2014; Benítez-Fernández et al. 2010).

The Language Issues

When it turned out that the Moroccan guest workers stayed longer in the Netherlands than expected, together with their spouses and children, the question posed itself as to what to do with the language background of those children. Obviously, the children went to school and were to learn Dutch, which they did. But would acquiring the Dutch language and culture imply losing their own language and culture? Shouldn't these have to be maintained or preserved, also against the background of the then idea that these children would eventually return to Morocco (which they did not). Moroccan initiatives started Arabic lessons – in a very rudimentary form. Later, in the mid-1980s, the Dutch educational authorities became aware of the question of what to do with the mother tongues of migrant children. It concerned Moroccan children, but also children with a Turkish background, the largest minority at the time (and still today), and children of various other ethnic origins. Eventually, an initiative called “Education in Original Language and Culture” was created, for a few hours a week, in primary schools. However, soon afterward a debate broke out on the nature of this “original language and culture.” Another point was, namely, that the Moroccan authorities interfered strongly in this debate, stating and propagating Standard Arabic as the language to be taught. But what use would that language be to children who would hardly use it in the Netherlands, and whose parents were by majority illiterate or semi-illiterate? Why not educate them in Darija or Berber to better communicate with them and reinforce their identity? The reply of the Moroccan authorities was that, in the first place, Standard Arabic was the one and only official language of the Kingdom – at the time – and, therefore, the language to be taught; secondly, and lastly, respectively, that Darija and Berber were hardly codified, and that Berber was divided into diverse varieties, and as such not teachable languages in their view.

Another complicating factor was that the Moroccan authorities insisted that they would supply teachers from Morocco for these lessons, something the Dutch side did

not want to allow – out of fear of political manipulation. In other European countries, national governments had no problem with the recruitment of teachers from Morocco, but the Netherlands was an exception to that rule. The issues were never really solved. In the first years, the lessons mainly consisted of Standard Arabic teaching, but at a certain point the nature of these lessons changed, and they were then called “Education in Allochthonous Living Languages (EALL),” the explicit goal being to teach the actual mother tongues of the children, i.e., Darija and Berber, and Kurdish for Kurdish speaking children from Turkey. Teaching materials were developed in the mother tongues, but at a certain point integration politics changed, drastically, in the Netherlands: from a multicultural, to an assimilative agenda (see above). Newcomers, or migrants were to become Dutch, use the Dutch language, and accept Dutch culture and values. The lessons of EALL were in the end suspended in 2002 and never came back. The author of this chapter has been very active in this debate and in this research, which was initialized and led by Guus Extra of Tilburg University, a tireless supporter of teaching migrant children their mother tongue (Extra and Yağmur 2004; more on the relevant debate at the time in Obdeijn and de Ruiter 1998). In a European Funded Comenius project, Tilburg University, together with colleagues from France, Germany, and the United Kingdom, developed teaching materials in Tarifit Berber, Darija, and Kurdish, but these were never actually used as such, also because the EALL lessons were at some point of time suspended (Aarts and de Ruiter 1998; Benjelloun et al. 2001; De Ruiter et al. 2009). The debate on the issue of the Berbers’ mother tongue never tuned down. It does not, however, play a significant role, anymore, in the Dutch educational system today. Berber associations, and leading men and women, treated in this chapter, are still active in promoting their language and culture, but many limit themselves to activities on social media, with little impact on the debate on the issues discussed.

The Berber Community in the Netherlands: A Case of Transnationalism?

Earlier research on the organizational structure of the Moroccan community in the Netherlands (Brouwer 2006) and the Amazigh community within it (Van Amersfoort and van Heelsum 2007) raises the question of whether the broader Moroccan community and Amazigh community could be regarded a case of transnationalism. This term relates to new forms of migration, as in the present times communication with the homeland is available, without ceasing, through the Internet and social media, in contrast to the not so distant past when contacts were established arduously through writing letters, if migrants were at all literate; or through rare and expensive phone calls, cassette tapes, or just oral messages to the family via friends going back to visit the home country (Vertovec 1999, 2001). Brouwer (2006: 1154) quotes Portes, Guarnizo & Landolt (1999), who describe the concept of transnationalism in terms of “occupations and activities that require regular and sustained social contacts over time across national borders” (1999: 221). In her study on the use of websites

by the Moroccan community, Brouwer focuses on two well-visited websites at the time (2006): *maroc.nl* and *Maghreb.nl*, still in existence. Both websites can be considered relatively loyal to the Moroccan government and monarchy; although, they do not avoid criticism, particularly of the government. *Maghreb.nl* has as its motto “The virtual house of Morocco in the Netherlands,” and *maroc.nl* is a forum where members, not necessarily Moroccans, express their opinions on a wide variety of subjects, often related to Morocco. The site contains some Islamic discussion groups but one, labeled “Imazighen,” has only 39 members (as of August 14, 2020). The site gives the impression that is favorable to the homeland, Morocco. *Amazigh.nl* has been an alternative forum for those who, exclusively, want to deal with Amazigh issues. The forum has more than 9,000 members. Whatever the case may be and returning to the question of whether the Moroccan and Amazigh community should be considered a case of transnationalism, it is remarkable to see that both studies are quite cautious in their conclusions. Brouwer, in the last sentences of her article (2006: 1167), states: “These types of websites can therefore be understood as a symbolic space that is part of the Dutch Moroccan community and which expresses their collective identity. In fact, what these websites keep together is not the transnational but the national network of Dutch Moroccan youths.” Van Amersfoort and Van Heelsum (2007: 259) conclude: “It is perhaps premature to speak of a ‘transnational Berber community’, but international developments will certainly have an impact on the relationship between Dutch society and its Moroccan immigrants.” It is clear that the wider Moroccan community, and the smaller Amazigh community, are in constant contact with their homeland, Morocco. Meanwhile that does not imply that all Moroccans are constantly focusing their attention on what happens “there,” as after all they have their lives “here.” Based on the developments, as described in this chapter, I would join the cautious observation of Van Amersfoort and Van Heelsum that labeling the Dutch Moroccan community as a case of transnationalism is, for the moment, still a bridge too far, and although Internet and social media will probably only grow in importance in the future that does not mean that transnationalism will grow in importance as well, because, after all, the whole world is, in one way or another, getting smaller, making the distinction of what is a case of transnationalism, and what is not, only more complex.

Outlook

What does the data gathered in this essay tell us? Of course, I first must state that what I entrusted to this chapter is my choice. I cannot possibly give a full account of all aspects of the Amazigh community in the Netherlands and the direction(s) in which it is moving. But some observations can be made. First: the Berber identity plays an important role in the various Amazigh organizations. They explore it, discuss it, value it highly, and they do their best to keep these languages and cultures

alive, not only for themselves, but also for the generations to come. The number of individuals with a Berber background, active in the public, or internal Moroccan debate, continues to grow due to outspokenness, and with a growing impact on the public opinion. In the heat of the strife of populists, and orthodox Muslims, Berber Moroccans try to voice their opinions on what it's like to live as an ordinary Berber, but also, more generally, as a person with "Moroccan" roots in the Low Countries. From that perspective one sees that the shift in attention that Van Amersfoort and Van Heelsom expected in their 2007 publication on Berber organizations from "country of origin" to "own identity" seems to be lagging or has even stopped. Van Amersfoort and van Heelsom (2007: 234) conclude that the political orientation toward Morocco has lost some of its importance in the course of generations and that "maintaining a core Berber identity among the immigrants has become the center of the organizations' activities in the Netherlands." Initially, Amazigh organizations in Holland had a political orientation toward the country of origin, motivated by the political circumstances of oppression that existed in Morocco at the time; and as time passed, they became more oriented toward questions of identity and the like. Although I agree with the observation of Van Amersfoort and Van Heelsom that Berber organizations are very much focusing on their identity, past and future, the overview I presented of the diverse Berber organizations gives rise to an expected continuation, or even a reemergence, of the country of origin as the focal point.

The Amazigh community, in all its diversity, tries to voice its Amazigh identity in the various debates that take place in the Netherlands, be it on dual nationality, the Hirak movement, or the glorious past of the Berbers in Morocco. It is, though, more difficult to determine whether the "Amazighness" of the Berber community in the Netherlands is getting through to the Dutch public. There is just no data on that subject. Another point to note is that Berber identity does not often appear as an independent subject in the public debate. It is often mentioned in a broader context when, for example, there is a debate about the democratic stance of the Moroccan government, or the issue of dual nationality. Nevertheless, the Berber identity is not completely hidden from the public: it plays a role in the Moroccan community, and from there it shows itself, from time to time, to the Dutch public.

It is to be expected that as long as the Moroccan authorities do not succeed, in the eyes of many Dutch Moroccans, in developing their country into a free and prosperous nation for all Moroccans, and continue to try to maintain bonds with Moroccans abroad and influence them, Moroccans in the Netherlands, and Imazighen in particular, will also continue to occupy themselves with this often unwanted influence and internal politics. A healthy democracy benefits from open, and sometimes fierce, debates. Only when voices are silent and no one dares to speak anymore, democracy is in danger. Fortunately, things haven't reached that point yet, and if all participants in the Dutch democracy, people with all possible backgrounds, Moroccan and non-Moroccan, keep an eye on their own, and the national interests, democracy is not in danger, no matter how persistent the populist and religiously inspired extremist forces threaten it.

References

- Aarts R, de Ruiter JJ (1998) Handleiding (in het Duits, Engels, Spaans, Frans, Koerdisch, Nederlands, Tarifit, Turks en Arabisch) bij de Serie 'Mijn eigen Taal'. Syntax Datura, Tilburg
- Abbtöy (2008) Cartoons van een Berber. Xtra uitgeverij, Amsterdam
- Abbtöy (2011) De s van het kwaad. N.P
- Aynan A (2020) Een erwt maakt nog geen snert. Van Oorschot, Amsterdam
- Benítez-Fernandéz M, de Ruiter JJ, Tamer Y (2010) Développement du plurilinguisme. Le cas de la ville d'Agadir. Espaces discursifs. l'Harmattan, Paris
- Benjelloun S, Bos P, de Ruiter JJ (2001) Llugha dyâl-î, 3. Babylon/Dutch University Press, Tilburg/Oisterwijk
- Bos P, Fritschy W (eds) (2006) Morocco and The Netherlands. Society, economy, culture. VU University Press, Amsterdam
- Brouwer L (2006) Dutch Moroccan websites: a transnational imagery? J Ethn Migr Stud 32(7): 1153–1168
- Caubet D (2006) Shouf Shouf Hollanda: succesvol en Marokkaan. de Geus, Breda
- de Koning M (2012–2013) Between the prophet and paradise: the Salafi struggle in The Netherlands, Rev can d'études néerlandaises 33(2)/34(1):17–34
- de Ruiter JJ (1989) Young Moroccans in the Netherlands: an integral approach to their language situation and acquisition of Dutch. PhD thesis, Utrecht University
- de Ruiter JJ (ed) (2001) Multilingualism in Morocco and Moroccan communities in Europe. Lang Ling: Int J Ling 8
- de Ruiter JJ (2006) Les jeunes Marocains et leurs langues. Espaces Discursifs, 44. l'Harmattan, Paris
- de Ruiter JJ (2012a) De ideologie van de PVV. Het kwade goed en het goede kwaad. Rozenberg Publishers, Amsterdam
- de Ruiter JJ (2012b) The speck in your brother's eye. Islam's perceived war against the west. Rozenberg Publishers, Amsterdam
- de Ruiter JJ (2012c) (translation of Ibrahiem al Koeni). Goudstof. Atheneum, de Berberbibliotheek, Amsterdam
- de Ruiter JJ (2019) Morocco's languages: a sign of diversity? IDEE: Tijdschrift van het wetenschappelijk bureau van D66, (207)
- de Ruiter JJ, Ziamari K (2014) Le marché sociolinguistique contemporain du Maroc. Espaces Discursifs. l'Harmattan, Paris
- de Ruiter JJ, Saidi R, Spotti M (2009) Teaching minority languages: the case of Arabic. Plusvalorproject (www.plusvalor.eu), Milan
- Entzinger H (2006) Changing the rules while the game is on. From multiculturalism to assimilation in The Netherlands. In: Bodemann Y, Yurdakul G (eds) Migration, citizenship, ethnos: incorporation regimes in Germany, Western Europe and North America. Palgrave Macmillan, New York, pp 121–144
- Entzinger H (2014) The growing gap between facts and discourse on immigrant integration in The Netherlands. Identities 21(6):693–707
- Extra G, Yağmur K (eds) (2004) Urban multilingualism in Europe. Immigrant minority languages at home and school. Multilingual Matters, Clevedon
- Mourigh K (2003) A grammar of Ghomara Berber (North West Morocco). Berber Studies. Rüdiger Köppe Verlag, Cologne, p 45
- Mourigh K (2021) De gast uit het Rifgebergte. Cossee, Amsterdam
- Obdeijn H, de Ruiter JJ (eds) (1998) Le Maroc au coeur de l'Europe. L'Enseignement de la langue et culture d'origine (ELCO) aux élèves marocains dans cinq pays européens. Tilburg University Press, Syntax Datura, Tilburg
- Portes A, Guarnizo L, Landolt P (1999) The study of transnationalism: pitfalls and promise of an emergent research field. Ethn Racial Stud 22(2):217–237

- Ruiter JJ (ed) (2008) *Langues et cultures en contact: le cas des langues et cultures arabe et turque en France et aux Pays-Bas. Espaces discursifs*. l'Harmattan, Paris
- San M v (2018) Belgian and Dutch young men and women who joined ISIS: ethnographic research among the families they left behind. *Stud Conflict Terrorism* 41(1):39–58. <https://doi.org/10.1080/1057610X.2016.1266824>
- Slootman M (2018) The Dutch integration landscape. In: Slootman M (ed) *Ethnic identity, social mobility and the role of soulmates*. Springer, Singapore, pp 59–83. https://doi.org/10.1007/978-3-319-99596-0_4
- van Amersfoort H, van Heelsum A (2007) Moroccan Berber immigrants in The Netherlands, their associations and transnational ties: a quest for identity and recognition. *Immigr Minor* 25(3): 234–262
- Vertovec S (1999) Conceiving and researching transnationalism. *Ethn Racial Stud* 22(2):447–462
- Vertovec S (2001) Transnationalism and identity. *J Ethn Migr Stud* 27(4):573–582



Language Attitudes Toward Berber: Two Decades After the Official Recognition

10

Mohammed Errihani

Contents

Language of Morocco: Status and Functions	213
French	216
Berber	218
Measuring Language Attitudes	220
Anonymous Questionnaire About Language Issues in Morocco for a Book Project	221
Conclusion	225
References	226

Abstract

This study aims to reexamine the attitudes of Moroccans toward the Berber language two decades after it was first recognized – and later made official by the state of Morocco. The findings in this study are compared to research carried out in Morocco between 2005 and 2007, which was published in the *Journal of North African Studies* in 2008 (Mohammed Errihani (2008) Language attitudes and language use in Morocco: effects of attitudes on ‘Berber language policy’, *The Journal of North African Studies*, 13: 4, 411–428, DOI: 10.1080/13629380701800492.). The goal is to investigate whether these attitudes have changed with the officialization of Berber, or have remained unchanged. The conclusions reached are based on data collected in Morocco in 2018 and 2019 using various methods: participant observation, public comments collected from newspaper articles, and social media outlets dealing with the Berber question; also, a survey of 180 respondents. Two decades after the initial recognition of Berber and the call for its inclusion in the educational system, and a decade after it

M. Errihani (✉)
Purdue University Northwest, Hammond, IN, USA

gained official status in Morocco, attitudes toward this language appear to have generally become more positive, especially as a result of its officialization in 2011. However, although most respondents agree that Berber represents an essential component of Moroccan identity and heritage, many continue to oppose its teaching to ‘all’ Moroccans mainly on practical grounds, citing its lack of instrumental value to guarantee employment and upward mobility, the saturation of the linguistic scene in Morocco, and the politicization of the language policy that requires all schools to teach Berber, which the state has so far failed to implement.

Keywords

Berber · Language · Policy · Linguistic attitudes · Morocco

Multilingual polities are often characterized by inherent and complex linguistic and sociocultural problems, and whenever there is a language problem within a polity, no matter how long it is ignored, eventually a language policy is developed to address it (Fishman 1973; Schiffman 1996; Spolsky 2004). But the potential for implementing any language policy requires a certain congruity between the policy itself and people’s attitudes, beliefs, and practices. A language policy cannot succeed if it counters existing sociocultural factors. Therefore, in order for a language policy to be effective, a change in the status of the language is necessary, but more importantly, a change in attitudes and perceptions vis-à-vis the language, its culture, and its speakers is warranted: a process that essentially requires a genuine commitment of the state and policy makers to use all resources in order to bring about such a change.

The problem is that even if the state and policy makers were serious about changing the status of a language, changing people’s attitudes toward the language is no easy feat. Ruiz (1984) and Schiffman (1996) stress the symbolic nature of most language policies, and assert that there is always a discrepancy between the *de jure* policy and the *de facto* one – between the language policy as it is stated in official documents, and the actual policy, and how it is implemented on the ground. Such symbolic gestures become more pronounced, and often necessary, when language attitudes toward a minority language are negative, and, simultaneously, the state feels the pressure to do justice to the language in question.

Thus, language attitudes play a critical role in shaping and effecting language policies that could eventually lead to maintaining, promoting, reviving, and slowing down – or even preventing language shift and loss. This is especially true when people’s attitudes toward a language are positive, which often translates into positive attitudes toward the community that speaks the language in question (Fasold 1984; Spolsky 2004). The opposite is also true: negative attitudes toward a language make any attempts at promoting and changing its status a difficult task, and ultimately accelerate language shift, and eventually the language loss.

Language attitudes are dominated by powerful ideological positions that are largely based on the supposed existence of a standard, more desired, hence more powerful language, one that is viewed as having a higher status and garners admiration and approval. Milroy (2007: 137) argues that “prestige is conferred on language varieties by speakers, and speakers tend to confer prestige on usages that are considered to be those of the higher social classes.” Speakers are conditioned by these ideological positions, which lead them to believe that their attitudes toward a certain language are common sensical. However, prestige can be quite a slippery concept in that it is often closely connected with the powerful group that speaks the language and by extension is assigned to the group’s language.

As said, a language variety can be viewed negatively simply because its speakers happen to belong to a speech community whose societal station is lower social, wielding less power in that society. These factors are instrumental in making its variety deemed as substandard and inferior as opposed to the so-called standard and prestigious variety spoken by the social elite, or the more powerful. The stigma attached to the language variety is therefore automatically transferred to the group that speaks that variety, even though from a linguistic point of view there is nothing inherently positive or negative about a language, or variety (Labov 2006).

Language attitudes are not only connected to the concept of prestige and social attractiveness (Garrett 2007), as they are also associated with competence, and instrumental and practical benchmarks. Nowhere is this more noticeable than in Morocco, a multilingual setting where the linguistic geography is quite complex in that it consists of several languages and varieties in contact; not in competition in everyday interactions. As such, it is only natural that these languages generate different, and often conflicting, attitudes depending on their functions and status in the linguistic landscape of the country (Boukous 2008; Chakrani 2013; Marley 2004).

But before going any further in our discussion of language attitudes in Morocco, and the extent to which they have remained stable – or changed – in the course of the last two decades, a brief review of the main languages and the status and functions they serve in the Moroccan linguistic scene is warranted.

Language of Morocco: Status and Functions

Arabic is the official language of Morocco, but as of 2011, Berber (AKA Tamazight) has become another official language of Morocco. The question, however, becomes “which Arabic?” Is it Classical Arabic (al-lugha al-‘arabiya al-fusha, “the eloquent Arabic Language”), the language of the Qur’an and a large body of religious and literary texts, or is it its descendent, Modern Standard Arabic, the Arabic used in mass education and media?: or is it Colloquial Arabic (*Darija*), the vernacular spoken by Moroccans in everyday interactions? Even though the type of Arabic

implied as the official language is not lucidly made clear, it is assumed to be a hybrid of Classical Arabic and the so-called Modern Standard Arabic (MSA). The Moroccan Vernacular, *Darija*, on the other hand, is not acknowledged in the constitution, and only recently became a subject of debate when it was proposed as a potential language for teaching in pre-school and elementary education (csefrs.ma).

As such, there are at least three varieties of Arabic in Morocco. It is “la langue véhiculaire de la nouvelle civilisation qui s’installe dans le territoire correspondant aujourd’hui au Maroc” (It refers to the linguistic medium of the new civilization that settled in the territory currently known as Morocco) (Benitez et al. 2013: 17). Classical Arabic might enjoy a high status, but it is not tied to social class, per se; it is rather indicative of a high level of religious and literary scholarship, which does not necessarily translate into prestige, nor does it lead to upward mobility, especially when compared with French.

For those unfamiliar with Modern Standard Arabic (MSA) and its status in Morocco, one salient comparison to make would be to English: while Standard English and Received Pronunciation, for instance, can be acquired by an individual simply by virtue of belonging to a certain class in Britain, in the case of Arabic, belonging to a certain class has no bearing on the type of Arabic one speaks (Haeri 2003). In Morocco, the higher the social status, the less likely a high level of fluency can be expected in Arabic. In fact, upper class Moroccans tend to be less competent in MSA and more fluent in one or more foreign languages, namely, French. Moroccan Vernacular, *Darija*, remains the language of home and intimate settings, while it is considered to have a lower status than the other two varieties. *Darija* is also rather distinct from other vernacular varieties spoken in the Arab world, and in particular those of the Middle East, because of how heavily influenced by French, Spanish, and Berber it is (Benitez et al. 2013; Versteegh 2014).

Of all the three varieties of Arabic in Morocco, Classical Arabic is considered to be the most prestigious because of its association with the Qur’an and because it is regarded by most Arabs and Muslims as the language of truth, the archetype of linguistic purity, and the most logical language (Haeri 2003; Suleiman 1994). It is probably viewed as such because it is believed to have undergone little lexical or syntactical change throughout the centuries, although its phonology has been affected by the regional variations of the different Arabic dialects in the Arab world. In other words, while Arabic vernaculars have changed due to their contact with other languages (Eastern, Western, as well as indigenous), Classical Arabic has remained relatively intact, albeit its function as a medium of daily communication has ceased for centuries.

Although Classical Arabic is not the mother tongue of any Arabic speaker, it has not died in the same way Latin or Sanskrit has, nor is it expected to die anytime soon. It has in fact been well preserved in lieu of the many vernaculars adopted, as was the case with Latin. Haeri (2003) argues that, contrary to Classical Arabic:

Most classical languages have by now disappeared: Sanskrit came to be replaced by the local languages of India, and Latin eventually gave way to the European vernaculars, generation by generation, and domain by domain until even the Vatican stopped requiring it to be the language of prayers. But Classical Arabic as the language of the Qur'an continues to separate the sacred from the profane, writing from speaking, and prescribed religious rituals from personal communication with God. (p. 1)

Haeri goes on to claim that Classical Arabic has been preserved, in particular, thanks to the performance of daily prayers. But, in fact, there seems to be more to this preservation, especially when we realize that the majority of Muslims who are able to recite verses from the Qur'an for their daily prayers are generally illiterate and, in many cases, do not speak Arabic. There are other far more important factors that have made this preservation possible, such as the positive attitudes, admiration, and the respect that Arabic speakers continue to hold for this language as a potential unifying force for all Arabs and Muslims. Secondly, there is schooling and religious education. Classical Arabic has been preserved through schooling, especially through Islamic education, which includes the memorization and daily recitation of the Qur'an, the prophet's sayings, and Arabic poetry – including pre-Islamic poetry.

Modern Standard Arabic (MSA) is sometimes viewed as Classical Arabic adapted for modern times. It is a more contemporary, secular version of Classical Arabic, but without the formality and rigidity that generally characterize the syntax and lexicon of Classical Arabic. MSA is the variety used in government, law, education, and the media. It is also the medium of literacy, and oral and written communication used by the educated elites, and serves as a lingua franca across the Arab world (Benitez et al. 2013).

The third variety of Arabic in Morocco is Colloquial Arabic, or what is called *Darija*. It is the mother tongue of most Moroccans, and serves as a lingua franca between Arabs and Berbers as well. *Darija* is not a written language; it is a spoken variety of Arabic ("Arabic" is used throughout this paper as a generic variety that stands for both Classical and Modern Standard Varieties.) that does not enjoy a high status and is seen by many Moroccans as "Arabe vulgaire" (Versteegh 2014: 7), a deviant form of Classical Arabic that enjoys little prestige. Benitez et al. (2013: 17) argue that « les différentes variétés d'arabe marocain ont été, jusqu'à très récemment, peu valorisées. »

Darija also includes several regional variations, which are marked mainly at the lexical and phonological levels, but these variations do not necessarily impede communication among Moroccans from these different geographical regions. Even speakers of "Hassaniya," a variety of *Darija* spoken in the Sahara, are still able to communicate with speakers of other *Darija* varieties in the far North of the country, thanks to the process of leveling and mutual accommodation (Versteegh 2014).

Darija is marked by extensive code switching and borrowing from French, especially among educated Moroccans, women in particular (Trudgill 1974), who

tend to use an amalgam of *Darija*, peppered with French words and phrases in their everyday interactions (Bentahila 1983). At times, this code switching will turn into triglossia and involves more than simply two languages. The majority of Moroccans will greet each other, for instance, by combining Classical Arabic, *Darija*, and French. Therefore, it is not unusual to hear *Assalamou alaikoum* ("Peace be on you" [Classical Arabic]), *Labas* ("Is everything well?" [Darija]), and *Ça va?* ("How is it going?" [French]). Responses to these greetings could also be in any combination of these languages. Code switching between *Darija* and French is quite common, and it is a sign of education, and higher social class and prestige to be able to juggle between these two languages with ease.

French

French has had a contentious history in Morocco and will probably continue to be a source of controversy for the foreseeable future. Attitudes toward French vary greatly because of its legacy of colonialism as many still see it as an extension of a colonial past during which it was the language of government and education in Morocco, during the French Protectorate (1912–1956). However, on the eve of Morocco's independence in 1956, one of the first priorities in the process of nation building was the institution of Arabic as the official language at the expense of French, thereby setting the tone for what's to come, namely, the Arabization policy. The state, led by the Istiqlāl Party (Party of Independence), the leading nationalist party in Morocco at the time, imposed a policy of Arabization on all institutions in an effort to promote the Arabic language and help it regain its lost prestige. Ever since, the Nationalist Movement, and the state, have pursued a policy of Arabizing all institutions, particularly the educational system, and this policy gained even more momentum in the 1970s and 1980s when nationalistic and Arabic ideals were at their peak (Pennell 2000; Marley 2004; Maddy-Weitzman 2011).

Nonetheless, despite the efforts of the policy of Arabization to eliminate French from the sociolinguistic scene in Morocco, it became obvious by the end of the 1980s that such an ideal was simply too difficult to achieve. Instead, decades of Arabizing the educational system have especially resulted in the creation of a new generation of monolinguals who did not have access to the same economic opportunities as their bilingual peers, who generally came from wealthy families that managed to bypass the policy of Arabization, and were able to send their children to private French or bilingual schools. In other words, the policy of Arabization short-changed the poor of Morocco, who were not able to pay for a French or bilingual education. Nowadays, the French language continues to fare well and is becoming an increasingly important language in all social and economic spheres, and Moroccans, especially those who were in favor of Arabization, seem to have finally come to terms with this bitter reality.

One might ask, “How did French come to be viewed as a dominant language on the Moroccan linguistic scene?” French has no official status in Morocco, primarily because of its controversial legacy of colonialism; furthermore, it is not acknowledged or mentioned in the constitution. Nonetheless, the state has unofficially declared this foreign language to be the language of business, finance, and more recently and officially, the language of science and technology in higher education, which makes it the *de facto* language of power and upward mobility. Therefore, it is only natural that it will continue to be perceived as the language of economic success, modernity, and progress – and the language of competition in the global market.

Bourdieu’s (1984) claim that it is the labor market that determines the value of any linguistic variety seems to apply very well in the Moroccan context. Arabic may be the official language, but it is by no means the dominant language. That title is reserved for French, and Arabic comes in second place. French is seen as a necessary tool to “maintain data bases, technical expertise, textbooks, and communication with nearby Europe. This is the final irony: the nation cannot function with the national language alone” (Berdichevsky 2004: 239). In an increasingly competitive world economy, the Arabic language seems to lack currency; it is deemed limited and requires the aid of French in order for the country to function in the twenty-first century.

Standard Arabic might be an important tool in the public sector, i.e., the market controlled by the state, but the private sector and the free market require not only French, but technical and scientific skills, which can only be acquired through French. Thus, in addition to its being Morocco’s most important tool for accessing the global market, French is also the medium of communication with the outside world. As such, it is viewed as the language of opportunity and openness toward the West. It is also what makes Morocco and its culture quite distinct from Arab countries in the Middle East. The average Middle-Easterner conceives of a Moroccan as a speaker of French and some dialect derived from Arabic, but not Arabic (This information is based on personal interviews and conversations during travels during time living in Egypt for one year (2011–2012) and travels in Jordan, Kuwait, UAE, Bahrain, Qatar, Saudi Arabia, Palestine, and Oman (2008, 2010, 2011, 2012, and 2017).).

Until the last few decades, Moroccans have largely been considered bilingual in French and Arabic as they are generally introduced to French at an early age, during their elementary education (third grade). This bilingualism has been on the decline as a result of the Arabization of the educational system that took place in the 1980s and 1990s. Nowadays, students who graduate from public high school in Morocco are not as proficient in French compared to those who graduated in the 1980s, not only because of a lack of qualified instructors in French, but most importantly because the scientific subjects that used to be taught in French are now being taught in Arabic in elementary and secondary education.

Tertiary education, especially its scientific subjects, has continued to be dominated by French, which is yet another indication of the failure of the process of Arabization in achieving its objective of jettisoning French on the educational scene in Morocco for Arabic (Pennell 2000; Marley 2004; Maddy-Weitzman 2011).

French used in Morocco is also marked by extensive code switching and borrowing,

a phenomenon that carries no stigma but is rather indicative of a certain level of education and status. It is known to be the preferred mode of communication among upper class Moroccans, a large majority of educated women, and business-people. It is also the preferred mode of written communication and social media. In short, French has increasingly become a necessary tool for anyone looking to compete on the job market; without competence in the French language, one is clearly condemned to low-paying jobs, or jobs in the public sector. Therefore, the majority of young Moroccan parents prefer to pay exorbitant amounts of money to send their children to French schools, or schools that emphasize French in their curricula.

French remains a controversial language, nonetheless, because of its legacy as the language of colonialism, and because of the competition it represents to Arabic, especially in the view of the nationalists, and the Islamists, and that is primarily why many political groups, namely, the Islamist party, and the PJD (Party of Justice and Development) still insist on the exclusive use of Arabic in all sociopolitical, economic, and cultural events. But the reality on the ground points to a different direction: Moroccans are embracing the French language like never before. One has only to turn on a local television channel to realize the importance of French in Morocco. The exclusive use of Arabic in public domains, the main goal of Arabization, has turned out to be a lost cause, and the Moroccan media has put the proverbial final nail in Arabization's coffin by making peace with French, bilingualism, borrowing, and code switching.

Berber

The Berber language spoken in Morocco consists of three main varieties: Tarifit (*Riffiya*) in the Rif Mountains region, Tashelhit (*Soussia*) in Southern Morocco, and Tamazight (*Shelha*) in the Middle Atlas region of Morocco. It is the language of communities native to North Africa and some parts of West Africa. Although Berber has been spoken in Morocco for thousands of years, it wasn't officially recognized as part of Morocco's linguistic and cultural heritage until 2001. This recognition came about as a result of years of continuous demands made by Berber activists who used language as the main platform, and core value, to demand official recognition of their linguistic, ethnic, and cultural legacies in Morocco. Initially, this recognition was primarily manifested in a top-down language policy

that required all school children to learn the Berber language, regardless of their linguistic or ethnic background.

The mandatory nature of this language policy that recognizes Berber by requiring everyone to learn it has given rise to a great deal of antagonism, and ill will toward the language itself, the government, the Berber activists, and IRCAM (Institute Royal de la Culture Amazigh), the institution set up by the King of Morocco to advise and promote the Berber language and culture. A 2008 study on language attitudes toward Berber, particularly toward the language policy that makes learning Berber mandatory for all school-aged students, found that the majority of Moroccans were opposed to the policy (Errihani 2008), especially its compulsory nature.

It is important to point out at this juncture that there exists no standardized Berber despite the fact that IRCAM has been working on standardizing the three varieties into one that could be taught in all schools. Thus, speaking of Berber, or the Berber language, could mean any of the three varieties mentioned above. The extent to which these three dialects are intelligible continues to be a source of controversy, even among Berber activists and researchers involved in the standardization efforts. The official discourse out of IRCAM is that the three varieties of Berber are mutually intelligible dialects of the same Berber language, which might be true when considering the concept of a dialect continuum. However, the average Berber speaker will admit that the three varieties can be quite distinct from each other, especially Tashelhit and Tarifit, due to the geographical distance between the two regions, where they are spoken.

IRCAM's insistence on the intelligibility of the three dialects seems to be politically motivated, and the case for standardizing all the three varieties is based on the argument that they are mutually intelligible and can be called a language instead of dialects. Referring to the Berber language as a cluster of dialects, instead of a language, might represent a potential hurdle for the maintenance and promotion of the Berber language, and would complicate the task of introducing it into the school curriculum. In the end, allowing three more languages into an already linguistically saturated scene would translate into more resistance on the part of the government, and the Ministry of Education, as it would present a new set of challenges to the educational system, which has been weathered by the policy of Arabization.

It seems that one way to overcome the three-dialect conundrum was the standardization attempt of all three dialects, a process that is still ongoing. However, standardizing these oral varieties into one that is intelligible across all regions of Morocco meant an alphabet was the first step in the process. However, that Berber activists were unable to agree: the French camp called for a Latin script while the more conservative activists opted for the Arabic script. Since a compromise could not be reached, it was agreed that a 5000-year-old script called Tifinagh, which is still used by the Touaregs of the African Sahel, could be adopted

as the script for Berber (Ameur 1994). The use of this script for all varieties of Berber was also meant to give the impression that there is only one variety of Berber that is intelligible across the country: a claim that is easily rejected by speakers of Berber themselves (Interviews of speakers of Berbers conducted over the last decade have consistently contradicted the assertion that all three varieties of Berber are mutually intelligible.).

The fact that Tifinagh is used as an alphabet that both teachers and students are introduced to, and have to learn, almost simultaneously, seems to be slowing – if not harming – the implementation of the policy of teaching Berber to all school children. The choice of Tifinagh was nothing but a political solution to a linguistic problem and is proof that internal strife and warring ideologies within IRCAM, and among Berber activists, have taken precedence over the interests of the students, the grass roots, as well as the success of this language policy. Choosing a foreign script as a new alphabet for Berber has had a negative effect on the attitudes of the average parents toward Berber, because they generally feel that their children are being used as pawns in the IRCAM game. Therefore, it should come as no surprise that the venture to implement the language policy of teaching Berber to all Moroccan children is lagging behind and continues to encounter obstacles prompted by policy makers, the government, and some teachers, too (Errihani 2017; Chakri 31 January 2020).

Measuring Language Attitudes

Language attitudes may be important to implementing language policy, but they are extremely difficult to measure because they are primarily veiled positions, and evaluative reactions, which tend to reflect people's opinions and stereotypes toward a language variety (Garrett 2007; Dragojevic 2018): “an attitude is an internal state of readiness, rather than an observable response” (Fasold 1984: 147), which makes any attempt at recognizing and interpreting attitudes toward language rather tricky, and at times problematic. Therefore, no single methodology is proven to be able to reveal language attitudes effectively and accurately (Fasold 1984). Thus, in order to investigate and uncover an accurate picture of Moroccans' attitudes toward Berber, two decades after this language has been recognized by the state of Morocco, several approaches were adopted to collect data on this study between 2018 and 2020, including: participant observation, public comments on newspaper articles and social media dealing with the Berber question, and a survey of 180 students from three universities across the country.

Below is the survey questionnaire used to collect data from university students. The approach used in this questionnaire is meant as an indirect method to elicit attitudes toward Berber mainly, although some questions may refer to other languages. The purpose is to limit acquiescence bias, thus, not to affect the validity of the research.

Anonymous Questionnaire About Language Issues in Morocco for a Book Project

1. I am Male ☐ Female ☐
 Arab ☐ Berber ☐ Both ☐ Not sure ☐

2. The official language(s) (الرسمية اللغات أو اللغة) in Morocco is / are: _____

3. Check one box: The language that should be used as a medium of instruction for scientific subjects in high schools and universities is
- a. ☐ Arabic
 - b. ☐ French
 - c. ☐ English
 - d. ☐ Arabic and French
 - e. ☐ Arabic and English

Please justify your choice:

4. Every student should be required to learn Berber (the Amazigh language) starting from elementary school.
- a. ☐ Yes
 - b. ☐ No

Please justify your choice:

5. Children should start school by being taught in their mother tongue (Darija or Berber)
- a. ☐ Yes
 - b. ☐ No

Please justify your choice:

Thank you for your time. If you have any additional comments, please add them here:

A summary of all the findings show that a representative sample of Moroccans appear to agree on the perceived status, prestige, and instrumental value of each of the four main languages (I consider Moroccan Arabic (Darija) to be a language

rather than just a variety of Classical or Modern Standard Arabic because of their syntactic, lexical, and phonological variance.) in Morocco:

MSA (Modern Standard Arabic): Perceived competence and intelligence tend to be attributed to anyone who is able to competently use MSA, especially in conversation. Although this variety is not used in daily conversations except in formal occasions, it does command a high level of stature and respect toward those who are able to use it competently when called upon. This is due to the fact that very few Moroccans are able to do so with ease because of the diglossia and code-switching phenomena that Moroccans are known for, and which tend to present a significant language interference that most Moroccans are unable to overcome.

Darija (Colloquial Arabic): *Darija* is marked by its perceived utilitarian function for daily communication; otherwise, it is stigmatized in comparison to MSA (*Darija* tends to be stigmatized only when compared to more standard and formal Arabic.). One interesting finding with regard to *Darija* is that close to 30% of the respondents to the survey questionnaire thought that it was the official language of Morocco. The only explanation one can advance is that the word Arabic remains quite ambiguous in Morocco as it can refer to any of the three main varieties discussed earlier.

French enjoys this perceived prestige, and high status, not only because of the power wielded by those who speak/use it, but also for being viewed as having an instrumental value in guaranteeing social and economic mobility. However, the data shows that Moroccans are generally divided when it comes to expressing their attitudes toward French. Those who are generally from the upper classes tend to view it favorably, first because they are most likely to be proficient in it, and secondly because they realize the advantages and social capital it confers on them. Middle- and lower-class groups, on the other hand, generally tend to express antagonistic views toward French, and by extension toward the elite who use it regularly. And because they are typically not very proficient in French, they tend to opt for Arabic and English as the two languages that ought to be encouraged and promoted as the languages of government, business, and education. This goes to show the far-reaching effect of Arabization on those who couldn't afford a French education, and who, in the end, have become painfully aware, not just of how disadvantaged they are as a result, but also of how large the gap that exists between them and upward mobility has become.

Another interesting finding from the survey questionnaire is that around 25% of the respondents indicated that French is an official language alongside Arabic. What this indicates is that for the average person who is not well-versed in government language policies, the most widely used languages in Morocco are indeed Arabic and French. However, the government never acknowledges that French has any status in the country; although, every educational reform insists on the promotion of foreign language learning (1999 National Charter of Education; 2009 Emergency Program; Strategic Vision Horizon 2015–2030).

Berber: The data collected for this project has shown that attitudes toward Berber have generally become more positive, particularly as a result of the official status it gained in 2011. Those opposed to the teaching of Berber to all Moroccans do it

mainly on practical grounds, often stating that while it is an important component of the Moroccan identity and heritage, it lacks the instrumental value to guarantee economic mobility. The small minority of respondents who were in favor of teaching and learning of Berber did so on the grounds that this language constitutes part of the Moroccan multilingual and multicultural identity, and as such it should be preserved and maintained by teaching it to future generations. How respondents identified themselves (whether Arab or Berber) did not seem to have an effect on their answers to questions about the need to teach/learn Berber.

The perceived stigma attached to Berber, in terms of social attractiveness and economic mobility, and the fact that it used to be connected with folklore and rural Morocco in the Moroccan psyche, seem to be attitudes that have abated thanks to the newly acquired official status. Even in social media and online platforms discourse, where people tend to be more forthcoming and less inhibited or concerned about voicing their opinions, there is a general acceptance of Berber as an integral part of Moroccan identity, albeit it continues to be viewed as unnecessary to learn for two main reasons: first, due to the saturated linguistic scene in Morocco, and second, and more importantly, because it is seen as impractical when it comes to education, international communication, and employment.

One of the respondents stated, “We should not be required to learn Berber because we are not going to use it in our studies. It’s necessary to learn other languages that can make our lives easier.” Another respondent expressed a similar concern by stating that “the main purpose of the educational system is to prepare students for the job market, and most companies are international, so we need to study subjects in a language that most of the world speaks.”

Here, both respondents seem to be concerned about the benefit of acquiring another language: to use in one’s education and to guarantee employment and economic mobility, none of which can be guaranteed by learning Berber, according to these subjects.

If “Attitudes towards language are often the reflection of attitudes towards members of various ethnic groups” (Fasold 1984: 148), then the tangible change in attitudes in the last two decades toward the Berber language during the last decade seems to translate into more positive attitudes toward speakers of Berber in the public sphere. What appears to have played a role in this change is the widespread presence of Berber in the media, and in television, in particular. One can conclude that the official status assigned to Berber in the 2011 constitution, which reiterates that Berber is a major component of Moroccan identity and culture, has played a major role in changing attitudes toward the Berber language and its speakers, although the officialization hasn’t truly resulted in more concrete steps that aim at promoting the language, at least according to Berber activists. One only has to listen to Ahmed Assid, the most vocal Berber activist in Morocco, express this feeling in almost every interview or speech he delivers: (<https://www.youtube.com/channel/UCHLKCyfceG8wBNWUiAGzxWA>).

The official status that Berber gained in 2011 was not necessarily done in good faith and appears to be more of a symbolic gesture meant to placate the Berber activists who have been calling for such status for decades. The symbolic nature of

this officialization (Ruiz 1984; Schiffman 1996; Spolsky 2004) is clearly felt not only in the sluggish implementation, but even in the way the so-called official status is stated in the constitution.

Upon close reading of article 5 of the text of the constitution, adopted by referendum on 1 July 2011, one can't help but notice the peculiar choice of words used to differentiate between the official status of Arabic and that of Berber. Article 5 opens as follows: "Arabic *remains the* official language of the State. The State works to protect and develop the Arabic language and promote its use. In the same way, the Amazigh language (Berber) constitutes an official language of the State, as it is a shared heritage of all Moroccans without exceptions" (the emphasis is mine) (www.maroc.ma).

The use of two words – *remains* and *the* – to refer to the official status of Arabic is intriguing to say the least, in that it appears to set the official status of Arabic apart from that of Berber. The fact that Arabic "*remains*" indicates that the addition of Berber as yet another official language will, neither affect, nor alter the weight and/or influence of the Arabic language on the Moroccan linguistic scene. The use of the definite article to refer to Arabic seems to imply that there is only one official language in the country. In other words, there is "*the* official language," Arabic, and then there is "*an* official language," Berber, and the distinction between the definite and indefinite articles should not go unnoticed here.

This does support the argument that the underlying goal of most language policies tends to be symbolic at best (Ruiz 1984; Schiffman 1996; Spolsky 2004). In other words, the goal of recognizing a minority language, or instituting it as a national, or official language, could be nothing more than a symbolic act often seen as politically necessary. Therefore, while the overt aspects of recognizing and officializing a language might be motivated high ideals, the covert policies typically show ulterior motivations.

Knowing that language attitudes toward Berber continue to influence and shape the educational and political landscape in Morocco, one still needs to ask two crucial questions:

1. Does literacy and the level of education play a role in attitudes toward Berber?
2. Does religion play a role in shaping attitudes toward Berber?

The answers to both questions seem to be yes, for on close analysis of the respondents' attitudes toward Berber, it is clear that literacy and religion play a large part in determining one's attitude toward the Berber language: the more educated and secular respondents tend to be, the more open and receptive they are to Berber and its teaching and learning. On the other hand, the more religious and less educated respondents are, the less enthusiastic and welcoming they are to officializing and promoting Berber since it would constitute competition for Arabic, which they consider to be superior to any other language as they believe it to be the language of truth and logic, and the archetype of linguistic purity (Haeri 2003; Suleiman 1994).

Conclusion

The goal of this study was to provide a glimpse of the attitudes of Moroccans toward the Berber language 10 years after it became an official language, and 20 years after this language and its culture have been recognized by the State as essential components of the Moroccan identity and cultural heritage. Earlier research conducted on language attitudes toward Berber on the eve of its first recognition (Errihani 2008) revealed that the majority of Moroccans were opposed to the inclusion of Berber in the educational system. Furthermore, this earlier research found that the majority of Moroccans defended the Arabic language and argued that it was the most useful in the country, which they did out of loyalty as part of asserting their Arabic and Muslim heritage, given that Arabic is typically linked with Islam. Moroccans felt it, as their duty, to speak in support of the Arabic language even while they implicitly might be in favor of being educated or having their children educated in French for economic reasons. Such “conflicting discourses transcend the individual and become part of a larger government discourse that openly pledges its support for Arabic, while implicitly adopting French as the language of the future and modernity” (Errihani 2008: 18).

Two decades after the initial recognition of Berber, and the call for its inclusion in the educational system, and a decade after it gained official status in Morocco, attitudes toward this language appear to have generally become more positive, especially as a result of its officialization in 2011. However, though most respondents agree that Berber represents an essential component of Moroccan identity and heritage, many continue to oppose its teaching to all Moroccans, and do so mainly on practical grounds, citing the saturation of the linguistic scene in Morocco, lack of instrumentality of Berber in guarantying employment and upward mobility, and the politicization of this language policy, which the state continues to support publicly but has so far failed to implement.

In other words, the government’s failure to implement a language policy instituted almost two decades ago has contributed to, if not fueled, the negative attitudes Moroccans show toward the inclusion of Berber in the educational system, as they continue to view it as a useless political tool that the government is using to appease Berber activists.

Thus, the instrumental value of language determines its socioeconomic importance on the linguistic scene, more so than its integrative value (Baker 2006). The integrative orientation of Berber into Moroccan society has improved with its officialization, which marks a positive change from attitudes toward this language prior to 2011. In other words, the Berber language and culture’s marginalization and exclusion from the public sphere in Morocco has become less pronounced with the recognition of the Berber language and culture as essential components of Moroccan identity. Nonetheless, the instrumental orientation of Berber continues to engender negative attitudes to the point that even Berber speakers themselves seem to be shifting toward Arabic and French, for they symbolize power and economic opportunity for their children. Such language shift toward majority

- Milroy J (2007) The ideology of the standard language. In: Llamas C, Mullany L, Stockwell P (eds) *The Routledge companion to sociolinguistics*. Routledge, New York, pp 133–139
- Pennell CR (2000) *Morocco since 1830: a history*. New York University Press, New York
- Phillipson R (2000) *Rights to language: equity, power, and education*. Lawrence Erlbaum Associates, Publishers, London
- Ruiz R (1984) Orientations in language planning. *NABE* 7:15–29
- Schiffman HF (1996) *Linguistic culture and language policy*. Routledge, New York
- Skutnabb-Kangas T (1999) Education of minorities. In: Fishman J (ed) *Handbook of language and ethnic identity*. Oxford University Press, New York, pp 42–55
- Spolsky B (2004) *Language Policy*. Cambridge University Press, Cambridge
- Suleiman Y (1994) *Arabic sociolinguistics: issues and perspectives*. Curzon, Richmond
- Trudgill P (1974) *Sociolinguistics: an introduction*. Penguin, London
- Versteegh K (2014) *The Arabic language*. Edinburgh University Press, Edinburgh

Word Boundaries in the Writing System of Standard Amazigh: Challenges from Tarifit Facebook Users

11

Naima Tahiri

Contents

Introduction	230
Languages, Scripts, and Writing Against the Background of the Standardization of Berber ...	233
The Word as a Linguistic Entity	237
The Noun in Tarifit	238
The Word Spelling in Standard Amazigh According to Boukhris et al. (2008)	241
Corpus and Method	243
Analyses	245
Definiteness Markers	245
Indefiniteness Marker	246
Prepositional Phrases	247
The Preposition <i>n</i> in Noun-Noun Constructions	248
Summary and Discussion	249
Conclusion	252
References	252

The content of this article corresponds to the paper published in German with the title *Herausforderungen und Perspektiven der Wortschreibung: Tarifit auf Facebook und Standard-Amazigh im Vergleich* (Tahiri 2022).

The author is a member of the “Laboratoire Sciences du Langage, Littératures, Arts, Communication et Histoire (SLLACH)” of the University Sidi Mohamed Ben Abdellah—Fès. This article was written within the framework of this membership.

N. Tahiri (✉)

Université Sidi Mohamed Ben Abdellah, Fès, Morocco

e-mail: naima.tahiri@usmba.ac.ma

Introduction

The present chapter attempts to take a relatively unconventional path of comparison: The word spellings applied by the speakers of Tarifit (a Berber language spoken in Northern Morocco and generally passed on by oral tradition) within an ethnically homogeneous Facebook group (one could also speak of an ethno-group) are compared with the orthographic rules for word spelling found in a reference grammar of Standard Berber (or Amazigh) conceived by Boukhris et al. (2008) (cf. the notes on the standard language and on the standardization of Berber in Morocco sections “Languages, Scripts and Writing Against the Background of the Standardization of Berber” and “The Word Spelling in Standard Amazigh According to Boukhris et al. (2008)”). The aim of the chapter is not an error-analytical examination of the posts and comments posted in the Facebook group, but rather a critical discussion of the rules of word spelling in nominal contexts established in Boukhris et al. (2008). Boukhris et al. (2008) claim to be the first grammar that does not deal with a specific variety or dialect of Berber, but one that is dedicated to the grammatical description of Standard Amazigh (SA). (In Morocco, Amazigh (or in the French spelling *Amazighe*; pronounced [amaziɣ]) has become established as the term for the standard language of Berber; in this chapter, the spelling *Amazigh* is used, which can also be found in English publications. Even if this term is to be criticized on a linguistic level, because *Amazigh* is basically a personal designation for a male person belonging to the Berber people, there is actually already a proper designation for the Berber language that includes all varieties or languages, which, however, is not used by Moroccan Berberologists. Moroccan Berberologists do not use this term because it is also used as a name for one of the three Berber languages or varieties in Morocco (*Tamazight*); the term *Amazigh* will be used in this chapter in particular to refer to the standard language of Berber developed by the *Institut Royal der la Culture Amazighe*. The corpus studied consists exclusively of linguistic data from Tarifit, the language or variety of Berber spoken in northern Morocco. The term *Tarifit* thus always refers to the language found in the corpus, while *Standard Amazigh* (SA) refers to the standard language discussed in Boukhris et al. (2008).) To date, this is the only reference grammar of SA. It was published by the *Institut Royal de la Culture Amazighe* (IRCAM). More recent grammars of SA have not been published by IRCAM to date.

The change in the status of Berber and its standardization process took its starting point in the royal speech of 17 October 2001 in Ajdir (cf. Royaume du Maroc 2001) and the subsequent foundation of IRCAM. The task of IRCAM was and is the preservation, the valorization as well as the promotion of the Berber language and culture. This was followed by the signing of various agreements between IRCAM and other state institutions to enable the integration of Berber in the public sphere. Two goals that have been achieved are the institutional implementation of Berber in the education system in 2003 and the establishment of a purely Berber radio station in 2010 (cf. Bouhjar 2012: 91). Despite all these efforts, Berber (both in the form of

SA and in the form of the regional languages or varieties) is neither offered in the school context comprehensively, nor is there broad social acceptance of SA or the Tifinagh writing system.

At least, however, the changed attitudes toward Morocco's Berber languages have had positive effects at the institutional level, reflected first and foremost in an amendment to the 5th article of the Moroccan constitution concerning the official languages. Berber is described as the cultural heritage of all Moroccans and is elevated to the status of an official language alongside Standard Arabic (cf. Royaume du Maroc 2011: 1766). However, its use in institutional contexts is almost nonexistent. Here, rather, an increasingly widespread use of French has recently become apparent, which also seems to be displacing Arabic. Although Arabic (meaning Standard Arabic) and Berber are Morocco's official languages under constitutional law, it can be observed in various institutions (including in the university sector) that French is preferred for communication by various parties. Taifouri (2022) even speaks of a "hidden 'Frenchification' of Moroccan education" (Taifouri 2022). The SA itself hardly receives any attention, which may have various reasons (more on this in section "[Languages, Scripts and Writing Against the Background of the Standardization of Berber](#)"). Nevertheless, Boukhris et al.'s (2008) reference grammar also has an iconic value. For a work that language teachers in educational institutions (should) also refer to, it is important that the orthographic specifications made in it are also linguistically tenable. The extent to which this applies in particular to word spellings in nominal contexts will be discussed below with reference to spellings by Tarifit speakers.

Tarifit speakers are organized in various virtual social networking groups or communities. Facebook seems to be a particularly popular platform for minority groups or communities that use a language without a significant written tradition for communicative purposes online. Using one's own language in written form becomes particularly difficult when the language in question is mainly orally transmitted. Where language users in social networks may be able to draw on the grapheme repertoire and knowledge of spelling rules of a related standard variety when writing in an uncoded language (as is the case, for example, with the use of Swiss German), the writing process is less problematic than with languages whose language users can only draw on grammatical and orthographical rules of languages that are not related or not closely related to the language in question. Such a situation exists in the case of Tarifit. Although Berber languages and Arabic belong to Afro-Asiatic, they belong to different language branches. Arabic is more closely related to Hebrew (via the Semitic language branch) than to Berber. There are major structural differences. Even though the Arabic dialect spoken in Morocco has converged with Berber on a structural and lexical level due to intense language contact, there is no mutual intelligibility between the speakers of the two languages. Like most Berber languages, Moroccan Arabic has no written tradition of its own. In the new media, however, an extensive use of this dialect in written form can be observed.

The lack of uniform, pedagogically taught spelling rules can lead native speakers of a language without a written tradition to use problem-solving strategies. The first hurdle to overcome is the choice of an appropriate writing system. The study by Tahiri (2016) showed that in Facebook groups, the Latin writing system is preferred for writing Tarifit, rather than the Arabic system or Tifinagh. (The Tifinagh [tɪfɪnɑːʃ] has its origin in the so-called Libyan inscriptions (datable to the sixth century BC). It is used in a modernized form (the neo-Tifinagh) in Morocco for SA. It is an alphabetic script consisting of unconnected characters, as seen in the following example: ⵜⴰⵎⴰⵏⵜ “house” (SA).) The next hurdle lies in the assignment of graphemes to sounds or phonemes (see also Tahiri 2016). Another challenge that must be overcome is the marking of word boundaries. Various decisions must also be made at the syntactic level, not least regarding the marking of sentence boundaries, for example.

In the present chapter, the focus is on word boundaries in the nominal context or in the context of nominal and prepositional phrases. A previous study on the setting of spatia in verbal contexts shows that Tarifit speakers tend to mark word boundaries relatively consistently when using their language within an ethnic group on Facebook (cf. Tahiri 2018). A comparison with the word spelling used in SA grammar (cf. Boukhris et al. 2008) shows a discrepancy between the spelling of laypeople and that of language experts. Boukhris et al. (2008) use spaces between verb stems and their surrounding morphemes more frequently than laypeople do (cf., for example, the separate spelling of affixes for the direct or indirect object in Boukhris et al. 2008). An investigation of word spelling in the context of nouns by Tarifit speakers in social networks will not only reveal the word concept of these language users, but will also serve in particular to take a critical look at the word boundaries set by linguists. Like verb stems, the lexical stems of nouns are surrounded by various morphemes that have a grammatical meaning and cannot stand freely in the sentence. However, there are also other grammatical morphemes that either precede or follow a noun and can be marked as free morphemes by spaces. This is the case, for example, with prepositions or postnominal units marking possession.

This chapter examines which nominal word boundaries Tarifit speakers set in the context of their written communication on Facebook. The aim is not only to show the similarities and differences between lay and linguist word spellings, but also to reflect critically on whether the current SA spelling rules—as laid down in Boukhris et al. (2008)—require revision.

Before the individual analyses or the comparison and critical comment can be carried out, the following sections deal with some background information or concepts relevant to the chapter, which, on the one hand, address the languages, scripts, and written form in the Moroccan context and, on the other hand, take up the concept of the word itself, the noun, or the nominal/prepositional phrase in Tarifit as well as the spelling rules established in Boukhris et al. (2008) in the nominal context.

Languages, Scripts, and Writing Against the Background of the Standardization of Berber

According to the 2014 census, 32.2% of the Moroccan population is illiterate. (See the website of the Haut-Commissariat au Plan: <http://rgphentableaux.hcp.ma/> (last accessed 18 September 2022).) On the other hand, the percentage of children between the ages of 7 and 12 who attend school is around 95%. The high proportion of illiteracy is therefore also a generational issue. Morocco's society is characterized by diglossia (Standard Arabic and Moroccan Arabic resp. Darija) and multilingualism (especially Arabic, Tarifit, Tamazight, Tashelhit, French, and Spanish), with Standard Arabic and Amazigh being the official languages mentioned by name in the 2011 constitution, while Moroccan Arabic and the three Berber languages Tarifit, Tamazight, and Tashelhit represent the first or mother tongues of the Arabophone and Berberophone ethnic groups, respectively. French enjoys a high prestige as a language of education nationwide, while Spanish is only regionally relevant. English is currently on the rise and—depending on the context—is competing with French in particular. Since the Skilled Workers Immigration Act of 2020 (“Fachkräfteeinwanderungsgesetz”), German has been one of the most popular foreign languages, the acquisition of which promises good educational and professional opportunities on the German labor market (cf. on the said Act <https://fachkraefteeinwanderungsgesetz.de/>, as well as on the status and acquisition of German in various institutions in Morocco Tahiri and Heming 2021).

Various writing systems are and have been used to write Moroccan languages in formal and informal contexts. Sometimes as early as kindergarten age (especially in private institutions), Moroccan children acquire the Arabic writing system in connection with Standard Arabic and the Latin writing system in connection with French and English. These are also the dominant writing systems that enjoy wide use and acceptance in Morocco. Focusing on the Berber languages of Morocco, the following is a brief overview of the challenges and perspectives related to the standardization process of Berber.

Language policy and the languages of Morocco have been a recurrent topic of discourse in recent decades, driven by different interests: Starting from the Arabization policy after the liberation from French colonialism since 1956 to the recognition of Berber as a cultural heritage and the attempt to give more value to Morocco's multilingual and multicultural reality. Even if the Berber languages are more or less more likely to belong to the orally transmitted languages and the official recognition as cultural heritage of all Moroccans is a relatively recent achievement, this does not mean that there was no writing or writing attempts in the past. For example, the Tifinagh is considered the first writing system with which various Berber languages came into contact, especially in the pre-Islamic period. This ancient writing system has been preserved by the Tuareg and is still used by them today. The Latin and Arabic writing systems were also in use, whereby the Latin scripts, which date back to missionaries, were or are used in particular by Berberologists, while the Arabic scripts were or are used primarily by clerics and poets in rural areas (Bouhjar 2012: 92; cf. also Pouessel 2008 for more detailed

information). Ould-Braham (2018) points out that the Berber manuscripts written in Arabic script, which were thought to be lost, have been discovered and collections expanded in the last three decades, but that most of them are almost inaccessible to today's researchers. At the same time, however, there are also freely accessible collections in public institutions (cf. Ould-Braham 2018: 191).

The old scripts or written documents must be considered marginal in that they do not or did not affect the Berberophone communities as a whole. However, they were the starting point for the choice of a writing system relevant to the standardization process. The decision was made in favor of Tifinagh, which was adapted to the needs of SA through various formal changes. The question of the script is also closely linked to the question of the standard. Against the background of a nonexistent homogeneity of Morocco's Berberophone language communities (however these may be defined or constituted), the question of the standard is not a simple one, because THE Berber exists neither in Morocco nor in other nation-states of North Africa:

Linguistically, [...] [Amazigh] is a linguistic entity fragmented into dialects and varieties between which the difference in morphological [sic!], lexical and syntactic structures may be as important as to render opaque the intercommunication between speakers from remote areas. The situation is much the same between the geolects of the North, Center and South of Morocco, although the difference has started to decrease with the development of media using the Amazigh language, the expansion of the teaching of the language and the organization of cultural and artistic activities in different cultural and linguistic areas grouping actors from different dialects. (Boukous 2014: 8; omission and addition N.T.)

Berber is understood in the present chapter as an umbrella term for all existing Berber languages/varieties and dialects, insofar as no explicit reference is made to a specific variety or language. There are more or less major differences between them on a linguistic level, which not only makes mutual understanding difficult but also impossible in nonadjacent language communities. This applies not only to Berber outside Morocco, but equally to Berber within Morocco. According to Boukous (2014: 9), the main problem faced in language planning is linguistic variation. IRCAM would have responded to this by taking geolectal reality into account. Competing forms would be treated as synonyms, and the choice of a standard form would be left to the user. However, Boukous (2014) also writes that "[...] in a second stage, a *normalized form* is favored, while leaving open the range of competing forms in order to raise the user's awareness to inherent variation in Amazigh as a whole" (Boukous 2014: 9; emphasis N.T.). Despite the variation-linguistic approach, the preference for a "normalized form" is advocated here in a further step. This approach can also be found in the reference grammar by Boukhris et al. (2008).

The introduction of Berber into the Moroccan education system in 2003 was to lead successively to an ever wider expansion of this new educational offer in primary schools. In her research on language teaching in public schools in Morocco, Pelligrini (2019) points out that in the three places where she conducted her field studies, according to statements by teachers and those in charge, very few primary

schools had implemented Amazigh teaching, and of the three schools where she conducted her studies, only one particular school offered such teaching. Pelligrini (2019) reports that one teacher in charge only came for half a day per week and taught the two 3rd grade classes in the school for 1.5 h per class. The lessons were characterized by cancellations and delays on the part of the teacher. These and other challenges related to teaching in SA are described in detail in Pelligrini (2019: 271). While the number of schools considered in Pelligrini (2019) is not representative and some time has elapsed since the study was conducted, it is reasonable to assume that these challenges persist across the country to this day, precisely because of the major impact of the COVID-19 pandemic on the entire education system. SA also tends to be a hurdle for the students, as it is very different from their own Berber variety or language, especially at the lexical level. In his study, Achamrah (2022) examined the attitudes of 20 teachers toward the teaching of Berber in Moroccan primary schools. Among other things, it was found that no practical benefit is attributed to Berber, that Standard Arabic and French are generally spoken in classrooms, and that only a small number of the interviewed teachers are able to read and write the Tifnagh script. The teaching of writing skills in Tifnagh is hardly taught in Moroccan schools (Cf. Achamrah 2022: 452). Critical voices of researchers on Tifnagh can also be found, such as Farhad (2017), who comments on the Tifnagh system in relation to the linguistic structures of Tarifit. Many milestones have been set in the standardization of Berber (especially in terms of legal frameworks), but much remains to be done to address the real needs of Berberophone communities and linguistic realities, and to create wider acceptance for SA teaching.

The use of Berber languages in social networks is unimpressed by language policy efforts. For example, Tarifit speakers communicate with each other in closed Facebook groups in their own language and use the writing systems available to them. The Latin writing system is preferred for writing Tarifit (cf. section “[Corpus and Method](#)”). The decision for a certain writing system or for certain orthographic rules within the writing practices of the group members should not be regarded as deviations from norms or conventions already existing or consolidated in the language community. As long as the norms prescribed by the institutions and linked to the SA are ignored by the Berberophone communities or no relevance is attributed to them for their own language and writing use, there are no conventions (such as those described in Sebba 2007: Chap. 2) for these communities that could be violated. What does exist is merely an in-group convention that enables group members to understand each other within the social networks. Deviations such as those described in Sebba (2007: Chap. 2) for the writing practices of speakers of written languages are therefore not the subject of the present chapter. In Sebba (2007), deliberately produced orthographic deviations from a normatively prescribed orthography are linked to social meaning. He speaks of “orthography as a ‘social practice’” (Sebba 2007: 26).

Even if there are no orthographic conventions or norms for the use of language and writing within the Facebook group studied in this chapter (cf. section “[Corpus and Method](#)” for more details) (at least against the background of the currently nonexistent or hardly relevant influence of the SA), it should not be disregarded that

the two written languages, Standard Arabic and French, were acquired at an early age. Although the influence of the SA on the speaking and writing habits of Berberophone communities is currently nonexistent or hardly relevant, it should not be disregarded that the early acquisition of the two written languages Standard Arabic and French, in Moroccan educational institutions, means that Tarifit speakers, just like all other Berberophone communities in Morocco, also bring with them written language competences and different forms of linguistic knowledge when they use their purely orally acquired first language in communication within the social networks. The types of linguistic knowledge mentioned in Mehlem's study (2007) entitled *The Graphematic Marking of Prepositional Phrases of Tarifit Berber in Written Narratives of Moroccan Students in Germany and Morocco* can also be partly assumed for the present chapter:

(1) the implicit knowledge about the oral structures of the first language, serving as a means of communication in family and informal community contexts, but not object of systematic phonological or grammatical reflection or metalinguistic awareness, (2) the explicit knowledge about the relevant script systems, Arabic and Latin, directly presented by formal instruction in school, concerning writing direction, letter shapes, letter names and other basic features of the script systems [...]. (Mehlem 2007: 196)

In addition to the implicit knowledge of the oral structures of the first language, which is not itself the subject of systematic reflection in terms of phonology or grammar nor metalinguistic awareness, there is also an explicit knowledge of the Arabic and Latin writing systems—meaning here the writing systems without reference to a single language—which are taught in school lessons. (This even goes so far that the speakers of orally transmitted languages believe that their language knows no grammatical rules and that everyone can speak as they wish, which can also be observed time and again among students of German studies in Morocco (cf., e.g., also the comments in Tahiri 2018: 511).) In addition to these two types of language knowledge, Mehlem (2007) mentions a third type: “(3) a knowledge about the writing systems of Standard Arabic, French and German that combines explicit knowledge of instruction (grammar and orthographic rules) and implicit knowledge which develops spontaneously in the process of learning to read and write” (Mehlem 2007: 196). Here Mehlem (2007) refers to knowledge about the writing systems related to the individual languages and thus also to knowledge about the grammatical and orthographic rules of institutionally taught languages. Since Mehlem (2007) also conducted research in Germany in addition to his research in Morocco, German is mentioned as well as Arabic and French. The explicit and implicit knowledge of German—but also of English—can be neglected for the present chapter. Language knowledge with reference to the grammatical and orthographic regularities of institutionally mediated languages in Morocco is certainly present in connection with Standard Arabic and French, but with reference to other languages (especially English, German, and Spanish) it is less likely to be present. Mehlem's study (2007) can only be used for comparison to a limited extent with regard to the results, which is due, among other things, to the age of the test subjects

taken into account (pupils from the 2nd to 10th grade), the place where the study was carried out (at schools) and the text forms produced. The majority of the texts produced by the pupils were written in Arabic characters (55 out of 56 texts), which also makes a comparison difficult, especially since in Arabic there are also unconnected letter forms, but also characters for short vowels do not necessarily have to be realized. Mehlem (2007) assumes that the factor for the almost exclusive choice of the Arabic writing system is the closer relationship of Berber to national traditions, which in his opinion precludes a writing system closely related to Western languages and civilizations (cf. Mehlem 2007: 208). For the Facebook group investigated in the present chapter, on the other hand, a clear dominance of the Latin writing system can be observed—just as in other Facebook groups with Tarifit as a means of communication.

The Word as a Linguistic Entity

Whether the word is thematized as a linguistic unit in a language, or whether it even exists as a concept with the same meaning as known from European languages, probably depends very much on whether the language in question is an orally transmitted language or one with a written tradition. There is an interesting note in a footnote in Dixon and Aikhenvald (2003) which concludes with a question: “It is likely that all languages with an established (non-ideographic) orthographic tradition do have a word for ‘word’.” Other languages tend to create such a term once they are exposed to writing. The interesting question is how many languages with no written tradition have a lexeme which corresponds to *word* in English, *mot* in French, etc.” (Dixon and Aikhenvald 2003: 3; emphasis in the original). A look at the Tarifit reveals that although there is the word *awar* (< *awal*), which can be translated, e.g., into English *word* or the German term *Wort*, other meanings can also be assigned that have something to do with speech, with what is said, or with speaking and the voice. The meaning “formal promise” is also covered by the term *awar* (e.g., *iwfas awar* “he gave her/him his word”). Such meanings are also found in German (e.g., *jemandem das Wort verbieten* “to forbid someone to speak,” *sich zu Wort melden* “to pipe out,” *jemandem sein Wort geben* “to give someone one’s word,” etc.). The extent to which the concept of the word and the underlying concept of a syntactically independent unit actually exists or is absent in linguistic communities without a written tradition cannot be clarified here. However, the following analyses will at least provide an insight into the word concept of the Rif Berbers.

People who are not familiar with linguistics are unlikely to think about the idea of what a word is or how word boundaries are to be defined if the language they use is a purely orally transmitted language. On the other hand, everyone has a more or less clear concept of a word if the language used is also a written language and there are rules for writing words. Of course, it is very important that the writing system and the orthographic rules of this language have also been acquired by the individual. A word or a form of a word is usually marked by spaces and thus also quickly

recognizable; at least this is true of today's European writing systems (cf. Haspelmath and Sims 2010). This is already one of the criteria for determining the linguistic unit word or word form: A word is distinguished from another word on the syntactic level on the basis of an orthographic criterion. A word is a sequence of letters which is distinguished from another sequence of letters by a blank space. However, as Haspelmath and Sims (2010) aptly formulate, the rules of word writing in languages with a long history of writing are to a certain extent traditionally shaped and the first time a language is written can lead to different opinions about where the appropriate word boundaries should be set: "The rules for orthographic word division are to some extent simply traditional in languages with a long written history. And when a language is first written down, language users often disagree on where to put spaces between words, and when a conventional spelling is agreed on, the decisions are sometimes clearly arbitrary" (Haspelmath and Sims 2010: 189). According to Haspelmath and Sims (2010), agreement on a certain spelling or orthography in such languages can sometimes be attributed to obviously random choices.

The orthographic criterion is not only problematic for nonwritten languages, but also for written languages, especially in cases where components of a word are sometimes written together and sometimes separately, depending on the context (as in the German example *er hat ihn **an**gesprochen* vs. *er sprach ihn **an***, in which the prefix *an* is either written together with the verb stem or separately).

In the research literature, further criteria can be found that help to differentiate from other linguistic units below or above the word boundary. Each criterion has its own weaknesses.

However, these will not be discussed here. Rather, it should be made clear that an unprepared and nonlinguistically based delimitation of words is not an easy endeavor.

The Noun in Tarifit

In Berber languages, the noun carries three grammatical markers: gender, number, and state. With examples from Tarifit, the three nominal categories will be dealt with in the following. It should be noted in advance, however, that these categories refer to both nouns and adjectives.

The gender, which is divided into a masculine and a feminine, is usually a category marked by the anlaut and/or auslaut. Masculines have an anlaut with a vowel sound, as is the case with *asennan* "horn," *ifri* "cave," or *ur* "heart." The vowel-initial sound tends to be the rule for masculines in comparison to the consonantal initial sound. In an interdialectal or intrasystem comparison, nouns with a consonantal anlaut can often be traced back to forms with a vocalic anlaut (cf., e.g., in Tarifit *fus* "hand" and *fuð* "knee," which are based on *afus* and *afuð* respectively). Other masculines contain an /r/ in the anlaut, which can be traced back linguistically to an /l/ (*reβhar* "sea" < Arab. al-bahr or *raz* "hunger" < laž). Only in

certain cases do feminines also have a vocalic anlaut (cf., e.g., *utjma* “my sister”). Feminines, on the other hand, are more characterized by the *t*-morpheme typical for Berber languages, which occurs in Tarifit either as a circumfix (*t*-. . .-*t*), a prefix (*t*-. . .), or a suffix (. . .-*t*). In Tarifit in particular, the fricative equivalents can also be found instead of these plosive dentals. On the subject of gender, cf. in more detail, e.g., Mourigh and Kossmann (2019: Sect. 3.1.2).

By means of these dental affixes, not only feminine forms can be formed, but also diminutives, abstracts, etc. can be derived. Basically, according to Kossmann (2007), the opposition of masculine and feminine gender in Berber languages serves to contrast certain meanings (cf. Kossmann 2007: 430). Thus, they stand for a contrast of “masculine” versus “feminine,” “large” versus “small,” and “collective” versus “singular.” The gender, however, remains feminine regardless of the semantic content of the individual word as such (cf. also Boukhris et al. 2008: 33–37). Gender and sex usually correspond in the case of humans and animals, although the feminine gender is also used for gender-neutral designations, as it is the case, for example, with generic designations in the case of animals or for the diminutive. However, nouns with a *t*-affix are not always derivational forms, and the dental morpheme does not necessarily have to be two-part, as in the case of native *θsa* “liver” and *θiri* “shadow” or the word *ryaβəθ* “forest” borrowed from Arabic. In other cases, the dental affix is also missing completely, such as in the loan words *ssimana* “week” or *nniβira* “refrigerator.”

The number category is also binary: singular and plural. Various processes are used to form plurals, and these are usually characterized by suffixation or the alternation of a root vowel with/without suffixation (cf. Mourigh and Kossmann 2019: Sect. 3.1.3).

The nominal category state is a category dependent on the syntactic context. Berber languages generally do not know any case markers. Instead, a distinction is made between a free state and a construct state which are in no way comparable to the case categories known, for example, from Indo-European languages. The free state represents the citation form of the noun and is found in particular syntactic contexts, such as it is the case of a subject in sentence-initial position (e.g., *argaz iffa ayrum* “the man ate bread”; free state in SVO position) compared to the position of the subject following the finite verb (*iffa wargaz ayrum* “the man ate bread”; construct state of *argaz* in VSO position). The construct state is marked also in other contexts, such as after the indefinite marker, after prepositions and other elements. The marking of the construct state is usually done in masculine nouns either by changing the initial stem vowel to /w/, or by *w*-prefixing (cf., e.g., the above example *argaz* vs. *wargaz* “the man”), in feminine nouns by the omission of the first stem vowel. However, there are also a number of nouns that are always realized in the free state, regardless of the syntactic context (cf. on the state category in Tarifit Mourigh and Kossmann 2019: Sect. 3.1.4, in SA Boukhris et al. 2008: 42–48).

In the immediate syntactic context of nouns—i.e., within a nominal or prepositional phrase—there may be other bound or free morphemes or words. In the following, the focus will be on certain elements that (can) immediately precede or

follow the noun. On the one hand, it is about morphemes or morpheme constructions that can be summarized in terms of their function as determiners. These include demonstrative/definiteness morphemes, indefiniteness markers (with the meaning “a” and “any/some”), and also possessive expressions. On the other hand, the prepositions occurring in prepositional phrases will also be discussed here, which, depending on the definiteness marker of the noun, can either directly precede it or be separated from it by a determiner.

With regard to the postnominal position, interest is focused, on the one hand, on the definiteness or demonstrative markers, which can be classified as bound grammatical morphemes. Their freely occurring counterparts are the demonstrative pronouns. The demonstrative pronouns, like the other free pronouns, can stand as a representative of the noun itself. The bound definite/demonstrative markers, on the other hand, can only follow a noun and cannot replace the noun itself, but only accompany it. Unlike the free demonstrative pronouns, the definite/demonstrative markers are not differentiated with regard to gender or number. A formal distinction is only made with regard to the proximity or distance to the speaker and/or listener. In Tarifit, these are *-a*, *-in* and *-nni*. In terms of their meaning, they can be classified both as reference-fixing definite markers of a noun and as demonstrative elements, as, for example, in *asrem-a ired* “the/this fish tastes good.”

On the other hand, possessive markers are also found in postnominal position, which will as well be considered in the present chapter. When considering possessive markings in Berber, a general distinction must first be made between nouns for kinship relations and the markings for the rest of the nouns. Possessive marking in most kinship nouns is usually done with the help of bound morphemes, as in *uma-f* “your (masc.) brother,” *uma-m* “your (fem.) brother,” and *uma-s* “his/her brother.” Only in the first person singular, there is no possessive suffix. An unmarked kinship expression always marks the first person singular possessive, such as *uma* “my brother.” For the other nouns and a few kinship expressions (e.g., *ajjaw* “grandson”/“nephew”), the possessive is marked by the preposition *n* “of” and a personal suffix. The possessive with the meaning “his” is then *n-əs*, for the meaning “our” *n-əx*, for “your (masc. pl.)” *n-wəm*, etc. Only the possessive form for the first person singular deviates from this: It is *inu*. Possessive expressions are only to be considered as companions if they directly follow the noun and are thus found within one and the same phrase. If, however, there are other elements between the two, then the possessive expression is not in the same phrase, as it is the case, for example, in the utterance *θaddarθ-a n-əs* “this house is his” (literally: *house-this his*; sense: “this house belongs to him”). In such a case, the whole construction is thus a sentence and not a phrase.

Indefiniteness markers precede the noun. There is no distinction according to gender. Indefiniteness occurs in Tarifit with the help of the numeral word *iɣ* “one” and the preposition *n* “of,” which in combination semantically correspond to the indefinite article *a*. The allomorphic form *ift*—which can be regarded as the result of assimilatory processes—only occurs before feminine nouns (*ift təmyarθ* “a woman”). Under the heading of indefiniteness markers, the gender neutral indefinite

pronoun *fa* or *fi* “any/some” preceding the noun is also considered at this point, which is used with entities that cannot be specified or are not known.

In addition to the abovementioned indefiniteness markers, prepositions, which in Berber also belong to the inflected parts of speech, can also occur immediately before a noun or between two nouns, which is why these will also be included in the following analyses. All prepositions can occur within a prepositional phrase (e.g., *zi Nador* “from Nador,” *ag uma* “with my brother,” etc.). However, in determinative constructions—i.e., constructions of the form noun-preposition-noun—only the preposition *n* “of” mentioned above can occur. The second noun determines the first (as, for example, in *θxadent n nuqarθ* “silver ring,” literally: *ring of silver*).

This list of various pre- and postnominal elements already outlines the focus of the corpus analysis. For a more intensive study of the grammar of Tarifit, see Mourigh and Kossmann (2019). For Berber languages in general and their genetic relationship, see Kossmann (2012).

The Word Spelling in Standard Amazigh According to Boukhris et al. (2008)

The grammar by Boukhris et al. (2008) is a descriptive and prescriptive grammar of SA. It is characterized by the use of two writing systems: the Latin system for the metalinguistic level held in French and the Tifinagh for the individual language examples. The book consists of a total of nine chapters. The second chapter, covering six pages, is devoted to spelling, which deals exclusively with the spelling of words or the use of spaces before/after certain morphemes, as well as the spelling of the schwa. Under the subtitle “Les classes de mot graphique,” there is a—partly seemingly arbitrary—list of 19 word classes to be marked with spaces: “nom” (specifically the noun), “nom de qualité” (meaning the adjective), “verbe,” “le participe,” “les pronoms objet direct et indirect,” “les pronoms autonomes,” and others. For the purposes of the present chapter, the focus will only be on the orthographic rules for writing nouns. In some cases, it is necessary to consider other chapters in Boukhris et al. (2008), as many aspects are not covered in the second chapter on spelling.

According to Boukhris et al. (2008), the noun is considered a lexical unit consisting of a root and a so-called schema (the vowel schema) and it can be primary, derived, or compound. For example, the noun *argaz* “man” is primary or simple, while a word like *buhijuf* “the hunger” is compound and one like *aslmd* “education/teaching” is derived (examples from Boukhris et al. 2008: 33, IPA notation N.T.). (Examples taken from Boukhris et al. (2008) and thus belonging to the SA are marked with “SA” in brackets in the following examples, as there are partly formal differences between the SA and the Tarifit. The IPA notations serve to make the examples readable for the general public. The object language level in Boukhris et al. (2008) is exclusively in Tifinagh.)

The noun carries—as already described above for Tarifit—three grammatical markers: gender, number, and state. For nouns (“nom”), which are not grouped together in a word class despite their formal similarities with adjectives (“nom de qualité”), Boukhris et al. (2008: 27) prescribe that the gender, number, and state markers be written together with the nominal stem, while determinative elements, such as the demonstrative or possessive markers, are written separately. The fact that gender, number, and state, as grammatical categories of the noun, should be written together with them is a logical conclusion from the highly obligatory nature and the high degree of abstractness of these inflectional categories and, in part, cannot be managed otherwise because of the stem-changing processes. From this point of view, the compound spelling of the individual morphemes into one word is therefore to be regarded as logical.

With regard to determiners, the situation is different. According to Boukhris et al. (2008), these should be written separately from the noun. This concerns the definite or demonstrative markers, such as in *abrid a* (SA) “this way” and *abrid inn / abrid ann* (SA) “that way,” or also the possessive markers, such as in *abrid nns* (SA) “his way.” Boukhris et al. (2008) also erroneously cite the example of *haqqa abrid* (SA) “that is the way” for separate spelling of determiners. Unlike the above examples, *haqqa abrid* is not a phrase structure, but rather a complete nominal sentence. Therefore, such an example should not be mentioned in the context of determiners. As an exception to the separate spelling of determinative elements, Boukhris et al. (2008) mention only possessive markers in kinship terms. In the context of (certain) kinship names, pronominal possessive affixes occur, which are then joined into one word according to the orthographic rule of Boukhris et al. (2008), as in *babas* (SA) “his father.” Only in the first person singular, there is a zero morpheme with pronominal possessive affixes, so that “my father” is then realized with the unmarked form *baba*.

Under Sect. 3.5.2 in Boukhris et al. (2008), the indefiniteness marker is taken up on p. 62. According to Boukhris et al. (2008), the indefiniteness of a nominal group is expressed by an indefiniteness morpheme. Mentioned here are the morphemes *jan*, *jun*, and *i33* (SA) for masculines and *jat*, *jut* and *ift* (SA) for feminines, as well as other morphemes expressing indefiniteness, partial relations, or uncertainty, such as *qra* and *fa* (SA). As with the indefiniteness markers in Tarifit mentioned above, the preposition *n* “of” is also used in SA. The examples given in Boukhris et al. (2008: 62) show the underlying spelling rules for word spelling: All morphemes are written separately.

According to Boukhris et al. (2008), prepositions are orthographically independent units that must be written separately from the noun that follows them, e.g., in *zi Midar* (SA) “from Midar.” However, the pronominal affixes of declined prepositions are written together with the preposition as one word. The preposition *n* “of,” which can stand between two nouns to form a determinative construction, is also separated by a blank space before the following but also after the preceding noun in Boukhris et al. (2008).

It is important to point out here that Boukhris et al. (2008) refer to the preposition *n* as “préposition génitive” in the context of determinative constructions and to noun-noun constructions with/without preposition *n* as “noms composés” (cf. Boukhris et al. 2008: Sect. 3.2.2). Thus, such noun-noun constructions are therefore also understood by Boukhris et al. (2008) as genitive constructions, which, however, must be regarded as nonexistent for Berber languages because the grammatical category of case does not exist. The transfer of grammatical categories and designations from Latin grammatical theory to languages like Berber leads to false categorizations that can obscure a clear view of grammatical phenomena. Not even French, which is closely related to Latin, knows genitive constructions. In French, there are no case markers in the nominal context. Where, for example, the genitive is used in Latin or German, French uses prepositions followed by formally unmarked nouns (cf. on the genitive problem in various languages, e.g., Glinz 1994: 108).

Corpus and Method

There are different Facebook groups by and for Tarifit speakers. The Facebook groups mentioned or studied in Tahiri (2016, 2018) can be assigned linguistically and culturally on the basis of the group name. Names of places, regions, cultural aspects, etc. are indications of the linguistic affiliation of the group members, as it is the case with group names such as *Ait Touzine* (name of a tribe in the Rif region) or *Nador inou* (city name and possessive pronoun in the 1st pers. sg.). The group name alone, however, is no guarantee that communication actually takes place in Tarifit. How different the choice of language or script can be in the groups was shown in Tahiri (2018) by comparing the three Facebook groups *Ait Touzine*, *tirifien waha* للرفيات فقط, and *IZraN NaZmaN*. The *IZraN NaZmaN* group stands out in comparison to the other two groups, especially due to a dominance of Tarifit (over 80% of the contributions or comments are written in Tarifit) and the Latin script (98.4%). It can be assumed that especially the thematic focus on a cultural aspect—as is the case with traditional rhyming songs (the *izran*)—promotes a higher frequency of Tarifit (cf. Tahiri 2018). Such Facebook groups are therefore more suitable for a study of Tarifit or Berber languages in general than those that are not formed around a cultural aspect.

The corpus used for the following analyses comes from the Facebook group *Izran ntahramin*, which was founded in June 2017. The name of the group means “rhyming songs of girls.” However, this does not mean that the group consists exclusively of female members. It is a private group that has more than 27,300 members at the time of data collection (October 2021). The following figure is a screenshot of the Facebook page of this group (cf. Fig. 11.1):

A private Facebook group is characterized by the fact that only group members can see the other members and their posts or comments. Nonmembers can find the group, but they must first be added as members in order to actively or passively



Fig. 11.1 Facebook group *Izran ntahramin*. (Screenshot: 14 October 2021)

follow the group events. The city of Nador is displayed as the founding location. Both the location and the name of the group are clear indicators of the expected language choice within the group. The corpus selected for the analyses was created by entering keywords in the search function of this Facebook group. The two terms “izri” “rhyming song” and “izran” “rhyming songs” were entered as keywords. Four posts were selected from a larger number of posts, which were followed by about 300 comments each. The difference between a post and the individual comments is that a post basically opens a new topic and the reactions to that post are called comments. The individual posts are from 2017, and the decision to use the search function to compile a suitable corpus is also due to the fact that such topic-centered Facebook groups move away from their topic focus over time. Currently, the posts within such a group like *Izran ntahramin* refer to very different topics and no longer exclusively self-formulated posts or comments are published. This fact makes it difficult to find adequate textual material for analyzing linguistic phenomena in Berber.

The number of collected posts or comments amounts to a total of 1245 units of varying length. Included are also those comments that consist only of nonlinguistic elements, such as emoticons. Also included are comments that contain not only Tarifit, but also other languages. For the writing of Tarifit, the Latin writing system is also preferred by these group members. Arabic characters in comments written in Tarifit are very rare. Since in Arabic certain graphemes always occur only in unconnected form, it is only logical to exclude from the analyses all units notated with the Arabic writing system. It would not always be obvious whether separate or compound spelling is intended. After sorting out all comments that do not contain

linguistic units relevant to the analyses or nominal contexts, a total of 956 comments (including posts) remain. These form the starting point for the analysis of writing behavior in previously defined nominal contexts of Tarifit.

Which grammatical units are analyzed in the nominal context is explained in detail above in section “[The Noun in Tarifit](#).” The individual analysis steps are subdivided as follows: definiteness markers (demonstrative markers and possessive markers), indefiniteness markers, prepositions in nominal or prepositional phrases, and noun-noun constructions.

Analyses

Definiteness Markers

Definiteness markings with demonstrative elements occur rather rarely in the corpus. Only 39 occurrences can be found. These are distributed over all three morphemes (-*nni*, -*in* and -*a*), whereby -*nni* is used somewhat more frequently ($n = 19$) than -*a* ($n = 17$) and -*in* ($n = 3$). The definiteness marker -*nni*, which marks the distance to the speaker and is also formally longer than the demonstrative -*a*, is written together with the preceding noun in 4 out of 19 cases (“*azini*” (the italics in the quoted corpus passages are not to be found in the original. They are used in the present chapter to emphasize certain forms) [*zzin-nni*] “this beauty,” ...). In most contexts ($n = 15$; 78.9%), the separate spelling is found (e.g., “*yemam ni*” “this (your) mother,” “*abrid ni*” “this way,” “*Atay ani*” “this tea,” and “*sawar ni nadjin*” “these jeans”). The morphemes -*a* and -*in*, on the other hand, are always written together with the noun (total $n = 20$; e.g., “*misaya*” “this table,” “*a3chiya*” “this afternoon,” “*nhara*” “this day,” and “*Yzriya*” “this song”). (Numerals are also used in the writing of Tarifit (but also in Moroccan Arabic) when the Latin writing system is used. This is because certain Berber (or Arabic) sounds that do not correspond to characters in the Latin system (such as the pharyngeal sound /ʕ/, which is represented by the numeral <3> in the abovementioned example) are represented by numerals (see also Tahiri 2016 for grapheme-phoneme mappings in Latinized transcriptions of Tarifit).) If we take all instances together, we see that compound spelling tends to be preferred (61.5%).

Possessive markings occur comparatively often in the corpus. As mentioned above, a distinction must be made between possessive markings for most kinship terms and those for all other nouns. In the context of kinship terms, only the markings for the first person singular, the second person singular masculine, the second person singular feminine, and the third person singular appear in the corpus. The first person singular in particular dominates with 87 of 168 occurrences. Since the first person singular is a zero morpheme (as mentioned above, the absence of a suffix indicates the meaning of the 1st pers. sg., as in *yamma* “my mother” and *baba* “my father”), the question of spelling with or without a blank space is obsolete here. The situation is different in the other contexts: 31 occurrences are in the second person singular masculine, 24 occurrences in the second person singular feminine,

and 26 in the third person singular masculine. In all cases, compound spelling is used exclusively, e.g., in “yemmasch” “your mother” (mas.), “babam” “your father” (fem.), and “Yamas” “his mother.” Among the kinship terms that occur in the corpus, the term *mother* dominates with 82 out of 168 cases, followed by *father* and *sister* (by mistake, in Tahiri (2022), the corresponding word for “daughter” instead of “sister” is written in this passage) (with 21 occurrences each). The terms *brother*, *siblings/brothers* or *sisters/brothers*, *daughter*, *son*, *grandfather/grandmother*, and *aunt* (on the father’s side) follow in descending order of frequency.

A total of 434 possessive markers can be found for nouns that do not belong to the kinship terms. The majority of the possessive markers are written separately from the preceding noun (82.5%; $n = 358$), and compound spelling can only be found in 17.5% of the cases ($n = 76$). The dominant possessive marker is the first person singular (*inu* “my/mine”) with a total of 249 occurrences.

While separate spelling is used with above-average frequency for all other possessive markers (in descending order of frequency, these are the markers for the 1st pers. pl., the 3rd pers. sg., the 2nd pers. sg. fem. and mas., 2nd pers. pl. mas. and fem., and finally the 3rd pers. pl.) (at least 91.4% and up to 100%), the marker for the first person singular is written separately from the noun in 72.7% of the cases.

There are two cases where the possessive markers for kinship terms and for other nouns also occur together: “awtachmayno” and “wtchma *inu*” each with the meaning “my sister.” The addition of the possessive marker *inu* to the already possessively marked noun *utfma* “sister” has an emphatic function here.

Indefiniteness Marker

With a number of 49 occurrences, indefiniteness markers belong to the rather rarely occurring elements in the corpus. The form *iz n* “a” is more frequent ($n = 32$) than the form *fa n* ($n = 17$). In most contexts, the indefinite marker *iz n* is written as one word and separated from the following noun by a blank space ($n = 23$), as for example in “*ijan rkas*” “a glass,” “*ijan wabrid*” “a way,” and “*ijan tabrat*” “a letter.” The writing of the indefinite marker with the following noun as one word follows in second place ($n = 6$). It is noticeable that compound spelling is used when the preposition *n* is not realized or when the anlaut of the noun contains an /n/, as in “*ijoraghi*” “a call” and “*ijanha*” “a day.” In two other cases, the anlaut of the noun is separated and joined to the indefinite marker: “*ijo raghi*” “a call,” “*ijo barad nwatay*” “a teapot.” In one case, only the preposition is attached to the following noun: “*ij natsawatch*” “a panty” (here in the sense of “inconspicuous trousers”). When using the indefinite *fa n* or *fi n*, the complete compound spelling with the noun dominates ($n = 8$), as, e.g., in “*xinizri*” “any/a song,” “*chinazin*” “any/a beauty,” whereby in one case the preposition is not realized (“*chowamchan*” “any/a place”). The second most common spelling is that of writing the individual components of the indefinite marker together while inserting a blank space before the noun, as, for example, in “*chan izri*” “any/a song,” whereby the preposition may also be missing (“*chi*.

wahboj” “any/a rude man”). Other spellings contain a blank space only between *fa* / *fi* and *n* (*n* = 2) or also between *n* and the noun (*n* = 1).

Although a more extensive corpus would have to be taken into account for a clear statement about the preferred spellings for *iz n* or *fi/fa n*, the available data show that putting a blank space before the noun with simultaneous compound spelling of the individual components of the indefinite marker tends to be the preferred option. Compound spelling of the indefinite marker with the noun is less frequent.

Prepositional Phrases

The spelling of prepositions that stand directly before a noun and do not carry personal affixes will be considered more closely in this section. The preposition *n*, which can stand between two nouns and has the function of establishing noun-noun constructions, is not considered here.

Only when this preposition is used outside of such a construction will it be included in the analyses. However, *n* rarely occurs outside of noun-noun constructions (*n* = 7). The most frequently occurring preposition is *ði* (with the allomorphs *ðəg/gil/g*) “in” (*n* = 147). The writing of prepositions together with the following noun dominates (60.6%), while separate spelling occurs in 39.4% of the contexts. In some instances, however, the word boundary between the preposition and the noun is not set correctly in the case of freely realized prepositions. Thus, sometimes the anlaut of the noun in particular is separated and connected to the preceding preposition, as in “*kho zedjif*” [x ʊzɛdʒɪf] “on the head,” “*ghal gharij*” [ʎa lxariʒ] “abroad,” “*jari limaniyan*” [ʒar ilimaniyən] “between Germans,” “*khal lahub*” [x lhub] “about love.” In a few cases, the final sound of the preposition is separated and connected to the following noun: “*gha roliman*” [ʎar uliman] “to Germany,” “*do ggharaf*” [dəg ʎaraf] “in the cup.” These are exceptional cases.

A closer look at the individual prepositions reveals that separate and compound spelling tends to depend on the formal length of the preposition. Phonetically short forms are more often written together with the following noun, and longer forms tend to be written separately. Thus, the prepositions *x*, *i*, *s*, and *n* are mostly written together with the noun (in 71.4–89.7% of the contexts). In contrast, the prepositions *sadu* “under,” *zar* (the right form is *zar* not *jar* as it is realized in Tahiri (2022: Sect. 7.3).) “between,” and *bra* “without” (11 cases in total) are always realized as free forms. The prepositions *ʎa(r)* “at,” *ag* “with,” and *zi* “from” are written together with the following noun in less than half of the contexts (42.2%, 37.5%, and 28.6%, respectively). The compound spelling of the preposition *ði* “in” with the noun (including the allomorphs of *ði*) amounts to 53.1%. Formally, the latter preposition is realized in different ways: in addition to “di”/“de,” also “dag”/“dhg”/“dak”/“dg” or “g”/“ig”/“ag,” “gi” as well as “d.” It is noticeable that especially the form “di” or “de” is more often realized with a blank space: 71.8% of the preposition *ði* “in” realized as “di” or “de” are not written together with the following noun.

The Preposition *n* in Noun-Noun Constructions

This section takes a closer look at noun-noun constructions which are connected by the insertion of the preposition *n* “of.” Depending on the context, the explicit realization of the preposition may be optional. There are a total of 281 noun-noun constructions in the corpus. The individual components are realized in 50.5% of the contexts in the following form: compound spelling of the preposition with the following determinative noun while inserting a blank space before the preposition ($n = 142$), as, for example, in “*ijan rkas n waty*” “a glass of tea,” “*amramrah nrabha*” “like the salt of the sea,” “*Mis narif ino*” “son of my Rif,” and “*afigha nsahra*” “O snake of the desert.” The preposition *n* is realized in some cases by adding a vowel before or after the /n/ as a syllable in its own right, as in “*atghadant anyamas*” “her/his mother’s ring,” “*Tamchonta anhafsa*” “this misfortune bringing Hafsa,” “*amis natboharacht*” “O son of a madwoman,” and “*Amchoma nalhob*” “this misfortune bringing love.” In the case of a noun beginning with /n/, there is no formal differentiation between the preposition and the anlaut of the noun: “*Tkhdant no9at*” “the silver ring” (actually [*θxaðent n nuqavθ*], literally “ring of silver”).

The second most frequent formal realization of the noun-noun constructions is that of the simple juxtaposition of the two nouns with the omission of the preposition *n*. A total of 33.1% ($n = 93$) of the noun-noun constructions have this structure. As a rule, the marking of the determinative noun is realized in the construct state, as, for example, in “*arkas ochafay*” “milk glass,” “*Ayajdid ojana*” “O bird of the sky,” “*girwast owabrid*” “in the middle of the road/path,” and “*atitawin odir*” “O grape eyes.” These are linguistically correct forms. However, there are also cases where the marking of the construct state is not or not adequately realized or where the omission of the preposition is ungrammatical, as, for example, in “*ayasis arif ino*” (the correct form for the italicized passage would be *n arif* or *narif*) “O daughters of my Rif” and “*go3rar ifosino*” (correct would be *ofosino*) “on the back of my hand.” A special case with a determinative noun is the separation of the anlaut (marked here in italics): “*bo yhakas i masma*” “the owner of shoes made of nails.” Reasons for the separation of the anlaut /i/ in “*i masma*” “nails” cannot be found here.

However, this spelling means that it is not possible to determine with certainty whether the commentator possibly classifies the /i/ as a preposition.

In the case of determinative nouns that do not require a construct state, a doubling would have to be realized in the anlaut of the noun if the preposition is omitted (the omission of the preposition in such cases is an assimilatory process that results in the doubling of the sound standing in the anlaut). However, this gemination is never marked by the writers, as, for example, in “*amis khatchi*” “O son of my aunt (on my mother’s side),” where the “kh” in “*khatchi*” would have to be geminated in the anlaut.

Complete separate spelling is less frequent in the noun-noun constructions. In 46 of the total of 281 contexts (16.4%), the individual components are each marked with a blank space, such as in “*Tadat n ra3do*” “the house of the enemy,” “*armoujat n rabha*” “O wave of the sea,” “*Ayamas n tasrit*” “O mother of the bride,” “*Tfawt n tatad nagh*” “the light of our house,” and “*ditawat n wakham*” “in the door of the

room.” This way of spelling testifies to high analytical skills of the writers. But also the fact that the individual nouns of the noun-noun constructions are never written together indicates a certain linguistic analytical competence. And where the preposition appears formally, it is either written together only with the determinative noun or realized separately from both nouns. However, the preferred spelling is clearly “X nY,” which is characterized by a blank space after the first noun.

Summary and Discussion

In the above analyses, the focus was on definiteness and indefiniteness markers, prepositions in prepositional phrases, and noun-noun constructions. In the case of definiteness markers with demonstrative suffixes, compound spelling dominates with 61.5%, whereby *-a* and *-in* are always realized in bound form and the formally longer *-nni*, on the other hand, is written together with the preceding noun only in about 21% of the cases. With 78.9% of freely realized forms of the demonstrative *-nni*, it is evident that the realization of the demonstrative markers as free morphemes seems to be strongly dependent on the phonetic length. In the above analyses, possessive markers were subdivided into kinship terms and all other terms. In the case of kinship terms, the possessive markers are always realized in bound form. In the context of all other nouns, there are also compound spelling of possessive markers with the preceding noun, but separate spelling clearly dominates with 82.5%.

Only few examples of indefiniteness markers were found in the corpus, so that the evidence of separate and compound spelling is not sufficient to be able to make unambiguous statements about the preferred spellings. But the insertion of a blank space between the indefiniteness marker and the noun with simultaneous compound spelling of the individual components of the indefiniteness marker tends to be chosen most often (59.2%). The majority of the components of the indefiniteness marker are thus realized together as one word. The phonetic brevity of the preposition *n* could play an important role in the frequent compound spelling. Where a preposition occurs, it is joined either with the preceding or with the following element, or with both.

The separate and compound spelling of prepositions in prepositional phrases can also be better understood against the background of the phonetic length or brevity of the prepositions. A total of 60.6% of prepositions are written together with the following noun, but the breakdown of prepositions according to phonetically short and long forms shows that the proportion of compound spelling is much lower for longer forms than for shorter forms. The factor influencing separate or compound spelling seems to be at the syllable structural level. An influence from the orthographic rules of Arabic cannot be ruled out here. In Standard Arabic, prepositions consisting of one grapheme (short vowel signs are disregarded) are written together with the following noun.

Also in the case of the preposition *n* “of” within noun-noun constructions, the Tarifit speakers tend to use compound spelling, whereby the connection of the

preposition with the following (the derminative) noun occurs most frequently (in half of the contexts), followed by the pure juxtaposition of the two nouns with the omission of the preposition (1/3 of the contexts). The nouns of these constructions are never written together.

With reference to the spelling rules formulated in Boukhris et al. (2008), it is advocated at this point to rethink the concept of the word or to mark it on a linguistically clearer basis and, in this context, to adjust certain specifications with regard to separate and compound spelling. The definite or demonstrative morphemes *-inn*, *-a* and *-in* or “ann” (SA), which according to Boukhris et al. (2008) must be separated from the preceding noun with a blank space, should be realized without a blank space for linguistic reasons. Boukhris et al. (2008) should establish the same rules for demonstrative markers as for possessive affixes in kinship terms, since demonstrative markers, just like possessive markers, are also bound morphemes or affixes from a purely linguistic point of view, which themselves require a free morpheme as a base morpheme in order to form a morpho-syntactic word. In both cases, they are nominal stems with which these affixes would have to be connected. Even if compound spelling for such units, which can be defined linguistically as dependent morphemes, is not always completely found within the Facebook group studied, the writers seem to have a good linguistic sense of the word concept overall. Clitics and affixes are similar in that they have a prosodic dependency, which is why they should also be linked to a prosodically independent element at the orthographic level. According to Haspelmath and Sims (2010), bound morphemes (which include clitics and affixes) are characterized by the following with regard to their boundness:

Clitics and affixes (collectively referred to as **bound forms**) are similar in that they exhibit **prosodic dependence**. This means that they cannot by themselves constitute a domain for word stress – they must ‘lean’ on a prosodic **host**. By contrast, canonical word-forms (often called **free forms**) exhibit prosodic independence. This can be seen in several ways. First, an utterance may be interrupted at a boundary between two free forms, but not at a boundary between a bound form and its host. This is true for affixes (e.g. *Paul ... started to play*, or *Paul started ... to play*, but not **Paul start ... ed to play*), and also for clitics [...]. (Haspelmath and Sims 2010: 196; emphasis in the original)

According to Haspelmath and Sims (2010), clitics and affixes show a prosodic dependency which manifests itself in the form that they themselves cannot be carriers of the word accent. Free morphemes or free word forms, on the other hand, are prosodically independent. According to Haspelmath and Sims (2010), this can be illustrated in various ways: An utterance can be interrupted at the boundary between two free word forms, but not at the boundary between a bound form and the stem (“host”), as exemplified in the examples mentioned in the quote above.

Even where it appears that bound morphemes carry an accent, this can be refuted by simple evidence. An example from Haspelmath and Sims (2010) can be used to illustrate this point: “In the French imperative *joue=le!* ‘play it!’, the weak object pronoun clitic *le* bears stress (*joue=‘le*), but this is the stress of the whole prosodic word (which happens to be on the final syllable), not *le*’s own stress” (Haspelmath

and Sims 2010: 196; italics in original). Thus, for bound morphemes that appear to be accented, this accent is to be classified as the word accent of the prosodic word. In the French word *joue-le* “play it,” the clitic element *le* does not receive an accent of its own but is only a carrier of the word accent, which here is on the last syllable. In general, when distinguishing between bound and free forms, free morphemes have more “freedom of movement” (Haspelmath and Sims 2010: 197) on the syntactic level. Bound morphemes cannot be moved on the syntactic level in the same way as free morphemes.

A suitable spelling for the definiteness markers *-nni*, *-a* and *-in* or “ann” (SA), regardless of whether they are considered by linguists as clitics or affixes, can only be compound spelling (as can also be found, for example, for the Tarifit in Mourigh and Kossmann (2019)). (In Mourigh and Kossmann (2019: Sect. 6.1), the definiteness or demonstrative markers are referred to as clitics (likewise in Tahiri 2022). However, the fact is that they cannot be replaced by their free counterparts in the same context. Therefore, it is more likely that they are not clitics but affixes. The statement in Tahiri (2022) that the definiteness/demonstrative markers are clitics must be corrected at this point.) The free counterparts are the demonstrative pronouns *wānni*, *wa* and *wīn* (for masculines) and *θānni*, *θa* and *θin* (for feminines), which can be used as proforms instead of nouns, can be shifted at the syntactic level and can also be considered as free words.

Possessive markers outside of kinship terms are also characterized by the linking of affixes (personal affixes) to a free morpheme (namely, the preposition *n*). In this case, Boukhris et al. (2008) prescribe the writing of affixes and the preposition in one word. In particular, the issue of compound spelling is not consistently thought through in Boukhris et al. (2008). For demonstrative markers, the same rule would have to be applied as for possessive affixes, since they also need a base morpheme to which they can attach.

If we look at the spellings used by Facebook users for demonstrative markers, for possessive affixes in the context of kinship terms and for possessive markers outside of kinship terms, we can say that in most cases the rule of attaching bound morphemes to a free morpheme is followed. In Boukhris et al. (2008), this rule is overridden for demonstrative markers. This needs to be reconsidered urgently.

For prepositions, Boukhris et al. (2008) prescribe separate spelling from the subsequent nouns. This also applies to the preposition *n* in noun-noun constructions. In the corpus of the present chapter, however, the opposite behavior can be observed for the most part: Depending on the phonetic length of the prepositions, either separate or compound spelling is preferred, whereby the amount of attached prepositions in prepositional phrases (over 60%) and in noun-noun constructions (around 50%) can be considered as relatively high. The compound spelling of prepositions with a noun should not be classified as an incorrect spelling option. In fact, bound adpositions can be found in different languages (cf., e.g., Dryer 2005). In Turkish, for example, certain postpositions are attached to the preceding noun as suffixes, as in the case of the postposition *ile*, which carries a comitative and instrumental meaning. The following passage from Lewis (2000) can exemplify

this fact: “It may be suffixed; the **i** is dropped after a consonant and becomes **y** after a vowel, the resulting **-le** or **-yle** being subject to vowel harmony: [...] **vapurla** ‘by boat’, [...] **dadiyla** ‘with the nursemaid’” (Lewis 2000: 84; emphasis in the original). Other formal adaptations mentioned in Lewis (2000) will not be discussed here. This brief example is merely intended to illustrate that bound prepositions should by no means be regarded as orthographically something aberrant (as mentioned above, monographical prepositions are realized bound in Standard Arabic) and that it is at least worth considering whether compound spelling should not be permitted, especially for the phonetically shorter forms of prepositions.

Conclusion

The exemplary discussion of the word spelling defined by Boukhris et al. (2008) is to be understood as a suggestion to deal more intensively with the orthography of SA. It cannot be assumed that the orthographic rules can be exhaustively covered in a chapter of only six pages, nor can the orthography of a language be dealt with solely from the aspect of word boundaries. The topic of orthography for SA requires more intensive reflection. Against the background of the results of the lay spelling of speakers of Tarifit examined here as an example, it could be shown that the specifications for word spelling made by Boukhris et al. (2008) do not necessarily coincide with the word concept of lay people, but also in part with linguistic theories.

The critical voices in the studies on the implementation of SA teaching in schools should also be taken seriously (cf. section “[Languages, Scripts and Writing Against the Background of the Standardization of Berber](#)” above). There are no adequate strategies for quality assurance of teaching. The standardization process has not yet shown any really relevant impact on Berberophone communities today. Research should be conducted into the factors that have a decisive influence on the widespread negative language attitudes toward SA and its integration in the public school system, in order to incorporate them in language policy decisions and the reform of educational curricula.

References

- Achamrah M (2022) Elementary teachers’ attitudes towards teaching of Tamazight language in Moroccan elementary schools. *World J Adv Res Rev* 13(02):448–459. <https://doi.org/10.30574/wjarr.2022.13.2.0043>. Last accessed 21 Sept 2022
- Bouhjar A (2012) Pour un aménagement linguistique de l’amazighe respectueux de la diversité géolectale. In: Agresti G, De Gioia M (eds) *L’enseignement des langues locales Institutions, méthodes, idéologies*. Actes des Quatrièmes Journées des Droits Linguistiques. Aracne editrice, Rom, pp 85–105
- Boukhris F, Boumalk A, El Moujahid EH, Souifi H (2008) *La nouvelle grammaire de l’amazighe*. Publications de l’Institut Royal de la Culture Amazighe, Rabat
- Boukous A (2014) The planning of standardizing Amazigh language: the Moroccan experience. *Iles d Imesli* 6:7–23. Online: <http://revue.ummto.dz/index.php/idi/article/view/356>. Last accessed 27 Jan 2022

- Dixon RMW, Aikhenvald AY (2003) Word: a typological framework. In: Dixon RMW, Aikhenvald AY (eds) *Word: a cross-linguistic typology*. Cambridge University Press, Cambridge, pp 1–41
- Dryer MS (2005) Order of adposition and noun phrase. In: Haspelmath M, Dryer MS, Gil D, Comrie B (eds) *The world atlas of language structures*. Oxford University Press, Oxford, pp 346–349
- Farhad EH (2017) Regard critique sur l’alphabet adopté pour enseigner l’amazighe au Maroc. Timsal N tamazight 8:19–35. Online: <https://cnplet.dz/images/revue/revue08/Regard-critique-sur-l-alphabet-adopt%C3%A9-pour-enseigner-l-amazighe-au-Maroc.pdf>. Last accessed 19 Sept 2022
- Glinz H (1994) *Grammatiken im Vergleich: Deutsch – Französisch – Englisch – Latein; Formen – Bedeutungen – Verstehen*. Max Niemeyer Verlag, Tübingen
- Haspelmath M, Sims AD (2010) *Understanding morphology*, 2nd edn. Hodder Education, London
- Kingdom of Morocco (2001) Dahir n° 1-01-299 du 29 rajab al khaïr 1422 (17 octobre 2001) portant création de l’Institut Royal de la Culture Amazighe. Bulletin officiel n° 4948 (1er novembre 2001), pp 1074–1076. http://www.sgg.gov.ma/BO/bo_fr/2001/bo_4948_fr.pdf. Last accessed 23 Jan 2022
- Kingdom of Morocco (2011) *Projet de la Constitution*. Bulletin officiel n° 5952 bis (17 juin 2011), pp 1765–1794. http://www.sgg.gov.ma/BO/bo_fr/2011/bo_5952-bis_fr.pdf. Last accessed 23 Jan 2022
- Kossmann M (2007) Berber morphology. In: Kaye AS (ed) *Morphologies of Asia and Africa*, vol 1. Eisenbrauns, Winona Lake, pp 429–446
- Kossmann M (2012) Berber. In: Frajzyngier Z, Shay E (eds) *The Afroasiatic languages*. Cambridge University Press, Cambridge, pp 18–101
- Lewis GL (2000) *Turkish grammar*, 2nd edn. Oxford University Press, Oxford/New York
- Mehlem U (2007) The graphematic marking of prepositional phrases of Tarifit Berber in written narratives of Moroccan students in Germany and Morocco. *J Written Lang Lit* 10(2):195–218
- Mourigh K, Kossmann M (2019) *An introduction to Tarifit Berber* (Nador, Morocco). Ugarit-Verlag, Münster
- Ould-Braham O (2018) Des manuscrits berbères anciens en graphie arabe. Quels objectifs ultimes attendre? *Faits Lang Soc* 3–4:190–211. Online: <https://revues.imist.ma/index.php/FLS/article/view/11520/6517>. Last accessed 20 Sept 2022
- Pelligrini C (2019) *L’enseignement des langues à l’école publique au Maroc: construction des savoirs, identités et citoyenneté*. Doctoral dissertation, Aix-Marseille. <https://www.theses.fr/2019AIXM0345>. Last accessed 19 Sept 2022
- Pouessel S (2008) Écrire la langue berbère au royaume de Mohamed VI. Les enjeux politiques et identitaires du tefnagh au Maroc. *Rev mondes musulmans Médit* 124:219–239. <http://remmm.revues.org/6029?&mp>. Last accessed 27 Jan 2022
- Sebba M (2007) *Spelling and society. The culture and politics of orthography around the world*. Cambridge University Press, Cambridge
- Tahiri N (2016) Zur Verschriftung mündlich tradiert Sprachen in sozialen Netzwerken: Der Gebrauch des Tarifit auf Facebook. *Glottology* 7(2):175–194
- Tahiri N (2018) About word segmentation when writing an orally-transmitted language: Tarifit speakers and their communication within a Facebook group. In: Boukhris F, El Moujahid EH, Bennis S (eds) *Langues et société au Maroc: Mélanges en hommage au Professeur Ahmed Boukous*. IRCAM et Faculté des Lettres et des Sciences Humaines Rabat, Rabat, pp 511–524
- Tahiri N (2022) Herausforderungen und Perspektiven der Wortschreibung: Tarifit auf Facebook und Standard-Amazigh im Vergleich. *Glottology* 13(2):111–142
- Tahiri N, Heming U (2021) Germanistik und Deutsch als Fremdsprache in Marokko. In: Tahiri N, Laasri M, El Mtouni S, Jai-Mansouri R (eds) *Germanistik und DaF in mehrsprachigen Kontexten*. Frank & Timme, Berlin, pp 43–70
- Taifouri M (2022) *Kalter Krieg um Marokkos Sprachen*. <https://de.qantara.de/inhalt/arabisch-franzoesisch-oder-englisch-kalter-krieg-um-marokkos-sprachen>. Last accessed 18 Sept 2022

The Berbers in Tunisia: Mistaken Identity Amid Contentious Politics in Post-Arab Spring Tunisia

12

Zouhir Gabsi

Contents

Introduction	256
Literature Review	258
Language Attitudes: Theoretical and Empirical Perspectives	260
Discussion: The Future of Berber Language in the Context of Language Contact	262
Language Contact and Attrition	263
Language Policy and Language Planning	264
Conclusion	267
References	268

Abstract

The Berbers of Tunisia form a small minority group inhabiting southern Tunisia with its rich and indelible history. They are ethnically and linguistically distinguishable from the rest of the population despite attempts by local governments, past and present, to describe them as *jbāliyye* “inhabitants of the mountains” with no reference to their linguistic and Berber heritage. It is a case of “mistaken identity.” They are not Arabs despite their bilingual situation, as there are no more Berber monolinguals in Tunisia. With Ben Ali’s fall in January 2011, there are renewed and emerging local and international voices that call to recognize this group as an ethnically and linguistically diverse group with new demands to teach the Berber language and hear it spoken in schools. However, these calls create a sense of division within the country, which drives the narrative that supporting the Berber cause may be regarded as treason. This debate is part and parcel of what is referred to as contentious politics. The chapter aims to discuss some of the issues related to the Berber case in Tunisia, and attempts to find solutions to safeguard this heritage without recourse to violence. In doing so, the chapter attempts to answer three fundamental questions: First, what is the Berbers’ current situation

Z. Gabsi (✉)

Humanities and Social Sciences, Deakin University, Geelong, VIC, Australia

e-mail: zouhir.gabsi@deakin.edu.au

today? Second, why do the Berbers feel crushed between their desire to maintain their language and culture – and to succumb to a separatist call ideology? Third, how can the Berbers in Tunisia restore their mistaken identity within the government language policy and demand to be recognized for their linguistic heritage?

Keywords

Tunisia · Berbers · Contentious politics · Amazigh · Douiret

Introduction

The Berbers in Tunisia represent a small minority ethnic group that inhabit southern Tunisia and the island of Djerba, with an approximate number of 0.5 percent (around 65,000 speakers) (Gabsi 2020). Despite this small number, there have been calls to recognize this group, especially after the Arab Spring in December 2010. This revolution brought significant changes, appreciated mainly in the freedom of expression and in the so-called “identity politics.”

With Berber, identity politics is understood to go beyond “sex,” “age,” or “race.” However, it seeks the “identarian logic where unity is sought beyond differences” (Lloyd 2005: 36). It is a usual occurrence that identity leads to politics: as stated by Barbara Smith “We have an identity, and therefore a politics” (Smith cited in Lloyd 2005: 36).

However, the question of Berber identity in Tunisia is far-fetched and overstated. This is confirmed by several fieldwork by Gabsi (2003, 2011, 2020). The lack of a persistent feeling of Berber identity in Tunisia is due to several reasons: demographic, geographic, economic, and political. Demographically, the Berber population is small and scattered in very isolated enclaves; therefore, the desire to unite and be vocal about their rights is undoubtedly doomed. This is worsened by the low socioeconomic level of these impoverished and marginalized communities.

Probably, the Tunisians, including Berbers’ main concern lies more in socioeconomics more than it does in sociolinguistics. Economic sustainability may be prioritized, as alluded to by Fishman (2001) in his *Can Threatened Languages Be Saved*. Though there is an awakened sense of Berber identity, and a desire to maintain Berber in post-Arab Spring, as bolstered by Gabsi (2020), these isolated communities’ socioeconomic situation subdues their voices to seek recognition. Anecdotal evidence through various fieldwork visits to the region suggests that the Berbers’ main concern is employment and the reversal of young Berbers’ migration to the major cities to further their studies and seek employment. These men’s departure causes a sociological *déchirure* (tearing) where villages become almost ghost towns, inhabited mainly by retirees and spinsters.

This sociological “tearing” is not the only problem the Berbers of Tunisia are facing. They have to confront the accusation of participating in the separatist ideology supported by local and international organizations. The Berbers feel sandwiched between the desire to maintain their language and transmit it to posterity

and the accusation that, by adhering to maintain their language, they are peeling off their Islamic heritage because Arabic is the language of the Qur'ān. As such, they inadvertently become embroiled in contentious politics. Contentious politics become part and parcel of the MENA (Middle East and North Africa) region, especially after the Arab Spring that swept the Arab world. It has corroborated, also, the idea that the Arab world has changed culturally and politically, and hence the "Arabs [and Berbers] belatedly regained their humanity and agency" (Gerges 2015: 1). This agency, through the *L'Association Tunisienne de la Culture Amazighe*, has awakened the Berbers of Tunisia to claim back their culture and reverse the mistaken identity from "the people of the mountain" to Berbers or Amazighs – and their language from "*dialectes Montagnards*" to the correct label of "Berber languages." According to Gabsi (2011, 2020), "*les dialectes montagnards*" is the name given to Berber languages in Tunisia as to covertly dismiss any Berber affiliation.

Inspired by the Berber movements and success in neighboring Algeria and Morocco, the Tunisian Berber Association has similar aspirations, despite its fledgling state. It aspires that Berbers have their language and culture recognized and have more access to the nation's services, and exercise more independence. In other words, these North-African states aspire to establish a unified land of *Tamazgha* "land of an imagined community" (Jay 2015: 340). However, even after a seemingly successful transition to democracy, the Tunisian government continues to disregard Berber as a separate language, which explains its absence in post-Arab Spring Tunisia's language policy. The lack of recognition has several reasons. First, there is a fear of the process of recognition, avoiding the violent clashes that occurred between the Arabs and Berbers in neighboring Algeria in what is often called the "Berber Spring" in 1980. In neighboring Morocco, Amazigh activists vehemently demanded recognition of their Berber language and culture, but resulted in police harassment and arrest (Silverstein 2012). However, after the Arab Spring that swept Tunisia and Egypt, there were mass demonstrations demanding the formation of a new Moroccan constitution with an emphasis on "regional, cultural, linguistic, and religious pluralism" (Silverstein 2012: 134). Second, the Tunisian government is more than familiar with the separatist ideology that drives Berber militants, and associations such as the Amazigh World Assembly, and the World Amazigh Congress. These associations fully support the Berbers in Algeria and Morocco – with little success – because they offer incentives to improve the situation in rural communities (Jay 2015: 341). Third, colonial history demonstrates that the French presence in North-Africa has used the separatist political stratagem to divide and rule. They treated the Berbers as though sharing more cultural affinities with France than they did with the Arabs (Jay 2015). During the French occupation period (1881–1956), the Berbers of Tunisia had to compete with the sedentary Arabs who heavily populated southern Tunisia. There were five groupings of Arabs and Berber concentrations in southern Tunisia with a grouping in Djerba, where autochthon Berbers, and Berbers of Beni Mzab, are found (Gabsi 2003). The ethnic separation between Arabs and Berbers was somewhat hazy during the French occupation, when the latter used this separation to gain more control over the country (Louis 1975: 31). In the early days of Tunisia's independence in 1956, only a few Berber villages were

fully aware of their ethnicity (Louis 1975: 31). Today, Berber militants, often the educated elite and Francophiles, including those in the diaspora, advocate a secular stance and aim to return to the pre-Arab presence in North Africa to better highlight and Berber identity and the sense of *Amazighité* (Jay 2015).

The preceding short introduction underpins the contentious political nature of the Berber cause in Tunisia. With the transnational mobilization of the Berber movement pertaining mainly to Algeria and Morocco, Tunisian Berbers followed suit and claimed recognition because of this “sudden” international interest in their cause and because of the unmissable opportunity to claim their true identity.

To understand the Berber question in Tunisia fully, the chapter proposes to answer questions about the present situation of Berber today and their attitudes toward their language and culture. The chapter covers almost a decade through various fieldwork in southern Tunisia outlined in this chapter’s methodological section. This is followed by discussing the legitimacy of the Tunisian Berber cause in a country riddled with economic mismanagement, marginalized communities, corruption, and political infighting.

Literature Review

The politicization of the Berber question is not a new narrative. The literature on the subject has taken the lion’s share of Berber studies mainly in Tunisia’s neighboring countries Algeria and Morocco, where the demographic variable plays a significant role in ethnolinguistic vitality. For instance, the number of Berber speakers is estimated to be around 25–30 percent in Algeria and 40 percent in Morocco (Silverstein 2012; Afrol News 2023).

Ethnolinguistic Vitality (EV) is a term developed by Giles et al. (1977: 186), who include the variables of “status,” “demography,” and “institutional support.” When comparing the EV of Tunisia, it scores much lower than its Algerian and Moroccan counterpart. This is primarily due to the low demographic number and the lack of institutional support. In Morocco, the creation of the Royal Institute of Amazigh Culture (IRCAM) aiming at integrating Tamazight into media and state education was met with dismay as Amazigh activists believe that IRCAM is only made of Berber mere folklore with no “real” political weight (Silverstein 2012). As for Algeria, the Algerian government body through HCA (*Haut Commissariat à l’Amazighité*), was created under the authority of ACALAN (African Academy of Languages) (ACALAN sponsors projects that harmonize languages in the African continent.), which was successful in meeting the membership criteria for cross-border lingua franca Tamazight (University World News 2019: para. 4).

However, the EV argument does not tell the whole story. In the Tunisian case, Berbers’ attitudes toward their language and culture are a significant determinant of its future. Recent work by Gabsi (2020) demonstrates that there is an awakened sense of Berber identity toward their language in post-Arab Spring Tunisia. It starkly contrasts with Gabsi’s work a decade ago, when informants have assumed negative

attitudes toward their language. For instance, he stated that an informant questioned “why are you studying our language? Our language is ugly” (Gabsi 2011: 135).

Though, understandably, one person’s attitude cannot speak for all Berbers, it, nonetheless, hints at a shift in attitudes: from a negative attitude, or possibly repressed voices, to claim their linguistic and cultural identity in pre-Arab Spring. This implies a positive consciousness of a Berber identity in post-Arab Spring Tunisia.

In pre-Arab Spring Tunisia, under Bourguiba the first President of Tunisia and his successor Ben Ali, the Berbers in Tunisia were not treated any differently from the rest of the population except for their language attitudes. The Berber language(s) are often classified as *les dialectes montagnards* – and the Berbers as *jbaliyye*, “the inhabitants of the mountains.” Gabsi (2020) explains this further by stating that this classification is the government’s ideology to present Tunisia as ethnically homogeneous, fearing that the distinction may risk its stability. Hence, research on Berber, any aspect of Berber, is frowned upon and treated as interference in Tunisia’s national affairs and sovereignty. Such concern is echoed in Gabsi (2003, 2011, 2020) and Battenburg (1999: 148). Battenburg states that:

Conducting research on Berbers in Tunisia is a daunting activity because one is questioning a fundamental tenet of the republic—linguistic and ethnic homogeneity. While the political environment precludes any type of empirical study about language use and attitudes among Berbers, analysis of the Berber’s role and status in Morocco and Algeria, along with a study of Berberphone communities in Tunisia, reveals the present state of this language.

In contrast, in post-Arab Spring Tunisia, and with the relative transition to democracy, the Berbers are enjoying the freedom of expression. Supported by local and transnational Berber communities, the Berbers in Tunisia have acquired renewed consciousness about their distinct identity and language. As discussed by Gabsi (2020), most Berbers, at least in the speech-Berber zone of Douiret, have strongly expressed their desire to maintain their language and get Berber the recognition it deserves as a vital language, vis-à-vis Arabic.

Such a view for the Berber language – to have a similar status as Arabic – is vehemently rejected by Labiadh (2017), a sociologist and previous Minister of Education in Ali Larayedh’s government. Labiadh argues that the Tunisian government’s attitude toward Berber does not emanate from marginalization as it was reported by Berber organizations, or Berber militants, but rather it is a question of its current use and future projection. For instance, the English language has gained new ground in the “linguistic market” as the preferred language of Tunisia’s present and future, due to its economic use, compared with French and Arabic (Sayahi 2014: 50) (See also Daoud 2011).

Labiadh adds that even the very concept of Berber, especially in southern Tunisia, became a marketing tool portraying Tunisia as a Berber country, where everything, such as artifacts becomes of Berber origin, mainly to attract tourists and foreign interest in the region. Despite Labiadh’s political background, his perspective remains a significant account of a Tunisian political discourse, illustrating and consolidating with evidence some of this chapter’s claims.

This précis of literature on the politicization of the Berber question in Tunisia marks wide gaps in research, especially on identity, language maintenance, and future work on the subject. Though efforts have been made on various aspects of the Berber sociolinguistic fabric of the Berbers from the 1950s with Basset's work, Brugnatelli (1987, 2000), and more recently with Gabsi (2003, 2011, 2020), we know almost nothing, or little, about the current state of Berber in terms of demography and language maintenance and shift in other speech-zones, such as Matmata, and the island of Djerba.

Language Attitudes: Theoretical and Empirical Perspectives

As mentioned earlier, EV (Ethnolinguistic Vitality) describes important factors that shape both status and future survival of a language, especially if it is a minority language with little institutional support. However, the concept of EV (approach) does not include behavioral and cognitive factors in the treatment of language maintenance and shift, as language is predominantly a social construct (Gabsi 2020). And behavioral and cognitive approaches hinge on the notion “that individuals are not born with attitudes but that they are learned, particularly over the course of socialisation during childhood and adolescence, although, in recent years, some researchers have propagated the notion that some attitudes may be inherited” (McKenzie 2010: 21).

From a socio-psychological theoretical perspective, attitudes mean perception, a filter that is maintained, or rejected, according to culture (McKenzie 2010: 22). This perception may be dictated by individuals or groups, consequently turning into “group perceptions,” or stereotypes.

The cognitive aspect of language attitude drives the speaker to value language in livelihood terms (Giles 1992; Garrett et al. 2003; Giles and Rakić 2014). In other words, the speaker assesses whether maintaining a language helps with the daily challenges, such as the speaker's financial situation. In the case of Berbers in Tunisia, most men leave their homes – the “safe house” for Berber survival – to larger cities where maintaining the language is a perennial challenge.

Furthermore, Garrett et al. (2003) have linked the cognitive with the affective variables in defining language attitudes and behavior. The affective aspect may indicate the Berbers' desire to hear their language spoken on the radio or television. Numerous Berber informants have expressed how fortunate their Berber friends in Algeria and Morocco are who enjoy the elevated status of Berber in those two countries.

Empirically, the data on language attitudes is based on several fieldwork visits focusing on the town of Douiret in 2003, 2014, and most recently in 2017. During the 2017 fieldwork visit, Gabsi (2020) surveyed 80 informants (46 females and 34 males) considering variables such as age, gender, place of birth, occupation, education, marital status, and family structure. Due to the paucity of the number of informants, only the variables of “age” and “education” will be used in this analysis. In the 2017 fieldwork, the survey comprised 23 questions split into two halves. The

first section covers questions (1–20) that concentrate on language fluency before and after the Arab Spring, where and from whom Berber learnt, and how the community felt about the language. The second section consists of three open-ended questions about the state of Berber in Tunisia, factors that support Berber, and barriers to Berber's language development. Both quantitative and qualitative approaches are used in this essay. The survey was conducted in person with the assistance of a local Douiri informant. Without her help, this study would not have been possible. However, even with this assistance, many informants were reluctant to provide extensive answers for possibly two reasons. First, it is either because informants, and Tunisians in general, are not used to this type of political activity, considered by them as intrusive, or because such actions may be due to "the shadow of the previous police state government is still affecting the psyche of many Tunisians, including the Berbers" (Gabsi 2020: 12).

Douiret, the subject of language attitudes study, is located approximately 9 km to the south of Tataouine, the administrative hub of southern Tunisia. Douiret comprises two parts: The Old Douiret "*Douiret el-qadīma*," and New Douiret or "*Douiret Al-jedīda*." Old Douiret is a cluster of run-down *qṣūr* (castles) with a population reduced to two or three Berber families. As in the case of many *qṣūr*, these ruined buildings are located on top of the mountain. The Old Douiret is located at a distance of 3 km from New Douiret. New Douiret is a small modern town, which offers modern commodities like running water, electricity, and telephones. Old Douiret was all but deserted by 1990 (Belhedi and Ouessar 1998).

Based on the 1984 census, the population of Douiret reached 1368 inhabitants. The average number in a family is six. The Census also included the number of rural families who preferred to live away from New Douiret. The number of rural families is 42, compared to 123 who have settled in the village. The number of rural dwellers is 297. It is interesting to note that Douiret's population in 1850 was 3500 inhabitants, nearly double its 1984 population. This shows that migration is an important factor in the town of Douiret. There are currently eight "tribes" (more correctly, families) in Douiret: El-ṣwābriyye, Awlād 'bīd, Awlād Ḥāmid, Awlād Buzīd, el-Zgādna, Awlād Belqāsem, Awlād Abdel-Karīm, and Awlād Ṭāleb.

According to Louis (1975: 50), Douiret is 700–800 years old. He claims they were the descendent of a Moroccan saint, "El-Ghāzi." He added that the immigration of Berbers from Algeria and Morocco under Almohads had considerable influence on the diversity of Berber dialects in Tunisia. A century ago, Douiri men lived off the produce of their gardens established on the side of small walls to get as much moisture as possible from the trickling water. They chose Arabs to look after their flocks, and at times they joined them to get some milk, or to bring back the animal herd home (Louis 1975).

Today, many Douiri men have either migrated to Tunisia's north or Europe. One informant, back in 2003, said that her husband prefers to live in Douiret, but the shortage of work is the only compelling reason for their migration. The Douiret inhabitants are welcoming people, especially when the visitor gains their trust; thus mirroring the local advice that when you visit a village, one should be in the company of one of its people (Gabsi 2003).

By and large, Douiret is a fortunate town compared to Chninni because the former has received more attention than the latter. Indeed, the establishment of the *Association de Sauvegarde de la Nature et de Protection de l'Environnement du Douiret* (ASNAPED) in 1986 has had an important influence on the town. ASNAPED aims to conserve the ecology, archaeology, and culture of Douiret. Dr. Belhedi, a fervent member of the association, believes that helping the Douiri through the promotion of eco-tourism and, consequently, the creation of independent local economy will stop the Douiri from migrating to other prosperous Tunisian cities. The association also believes that the Old Douiret *ghorfas* should be renovated so that some Douiri families, who initially abandoned it for a more comfortable dwelling in New Douiret, will find an incentive to make it their home again (Gabsi 2003).

According to Gabsi (2020), there is a renewed interest in preserving the Berber language in the town of Douiret. The age variable did not affect how the Berbers of Douiret see the symbiotic relationship between language and culture. For instance, 73 percent of those between the ages of 20-39, and 76 percent in the age group of 40-59 believe that the vitality of the Berber culture hinges on language survival (Gabsi 2020). As for the link between language and identity, most respondents have expressed, wholeheartedly, that language is a significant part of their identity. Interestingly, when informants were asked about the hurdles of Berber maintenance, most respondents point to the language difficulty as concerns pronunciation among young children. However, regardless of age or gender, most respondents emphasize the lack of the government's commitment to maintaining the Berber language, while they have expressed, vehemently, that Berber should have equal status as Arabic, and should be used on the radio and television since Berber was formerly the indigenous language of Tunisia (See Gabsi (2020) for a full analysis of the data.).

Discussion: The Future of Berber Language in the Context of Language Contact

The Berbers in Tunisia form an interesting case of study. They feel sandwiched between a strong desire to maintain their language, especially in the post-Arab Spring context, and the political consequences of such demands. The political nature of these demands is confounded by the Arab and Berber confrontation in neighboring Algeria and Morocco. The Tunisian government, past and present, know this rift, and is cautious of letting the Berbers of Tunisia have a say about the destiny of their language, or culture.

This "rift" is an exaggerated state of affairs in Tunisia. Transnational Berber associations have been driving a negative narrative describing the situations of the Berbers in Tunisia as oppressed and overtly marginalized communities. They support their arguments with far-fetched demographic published figures claiming that the Berbers in Tunisia represent 10 percent of the population, where in fact, the number of the Berbers of Tunisia does not exceed 0.5 percent (Gabsi 2020).

The small number of speakers, as stated earlier, has contributed to its questionable vitality. When coupled with other variables, such as the low socioeconomic status of

these Berber communities, one finds that the migration of Berber men to major cities in seeking work has contributed to language attrition, since the Berber men's migration usually results in migration of the family as a whole. Fishman's (2006) advice of "don't leave your language alone" could be reinterpreted as "don't take your language with you," as the only opportunity to speak Berber occurs mainly in the Berber speech zones; hence, its maintenance there far exceeds the options to maintain it elsewhere.

The Berber language maintenance in Tunisia has been, hitherto, the responsibility of mothers – they are unequivocally the custodians of the language. However, with the socioeconomic pressures on families many young mothers are forced to move to larger cities to seek work and be with their husbands. As a result, children born outside the Berber communities have little chance acquiring or retaining their Berber language. Unless economic development is planned for these marginalized areas, language attrition and eventual language death would not be a surprise.

Local Berber associations must initiate a discussion with the government to have a viable language planning, which includes the documentation of the remaining vernaculars, and by increasing its status by not advocating for it to have the same status as Arabic, but at least be recognized as a significant component of the Tunisian cultural heritage. Without institutional support, both materially and morally, the Berber language will face extinction in the coming decades.

Language Contact and Attrition

The Berber language has already shown signs of attrition, especially at the lexical level. Almost two decades ago, it was reported that 52.72 and 62.3 percent of the noun and adjective categories, respectively, are borrowings from Tunisian Arabic based on almost 1500 words (Gabsi 2003). The verb category seemed the least affected with 35.54 percent (Gabsi 2003). This is not surprising because verbs resist influence due to their syntactical significance: they inflect for person, number, gender, and tense (Gabsi 2003). Ennaji (1985: 13) confirms this by stating that

This rigidity [of verbs] may be accounted for that the verb nucleus is crucial for the syntax and semantics of the sentence, and it is at the level of the verb system that many morphological and syntactic processes are at work.

A similar percentage was also reported to be found in the Tarifiyt vernacular (Kossmann 2013) (Kossmann (2013) reported that, based on his corpus of 1500 words in Tarifiyt, 44% per cent are taken from the Arabic dialect.). In Tunisian Berber substituting native Berber words with their Arabic counterparts was usually observed in the speech of Berber children. For instance, words such as *tawqidit* (<Tunisian Arabic *wqīda*) has replaced the Berber word *tašlut* (Gabsi 2003: 231). Earlier Berber texts, particularly Motylinski's (1897), demonstrate additional examples of native Berber words on the island of Djerba that may have been lost. For instance, *taməsniau* "science knowledge"; *eml* "to show"; *təsəstoun* "interrogate";

tafaska “festival, celebration” (as cited in Gabsi 2003: 232). As expected with cases of language shift, Arabic loanwords in Tunisian Berber are present in many semantic fields, and they even contain lexemes that include numerals and body parts. This is a clear sign of the advanced shift toward Arabic in the Berber speech communities. The Berber language is not solely competing with Arabic and French, but also with the newcomer, the English language, due to its popularity on social media, especially among youth, i.e., its global economic hegemony.

Language Policy and Language Planning

Tunisia’s language planning and language policy are bound to its postcolonial history. After independence in 1956, Tunisia faced a linguistic dilemma, and had to choose between French, the language of Education and administration, and Modern Standard Arabic, the language of the Qur’an. Tunisian Arabic, the dialectal form of Arabic, was treated as uncouth, and it was hardly heard on local television stations compared to today, where advertisements are in Tunisian Arabic. Over three decades ago, Tunisian Arabic lacked prestige, and its status was comparable to what Berber is experiencing today. Maamouri (1983: 13) captured the low status of Tunisian Arabic:

It was generally thought that the “Tunisian dialect” was but a “degraded form” of the Arabic language and that it could not therefore interest any true scholars except for missionaries or Arabists whose aims were judged to be at least suspicious, if not condemnable (Maamouri 1983: 13).

Like the difficulties that Berber is facing in Tunisia now, Tunisian Arabic has been perceived by other Arabic speakers as a difficult language to understand, even though it is not evidenced by any research on mutual intelligibility, for instance. Equally, there is no research on perceived claims that dialects outside the Maghrebi dialects are considered closest to MSA.

The French language still plays a strong role in education and administration despite the onslaught of the English language. It would be important to ascertain language attitudes toward French in modern-day Tunisia. This is because for almost 40 years ago Ounali (1983: 110) observed that when he interviewed university students from the faculties of humanities, law, and science three-quarters of students preferred the French language to Literary Arabic, Modern Standard Arabic, and Tunisian Arabic.

Language planning in Tunisia had favored diglossia with Modern Standard Arabic and French as the choice languages of the nation (The concept of “diglossia,” refers to a situation where two distinct related languages are used side by side throughout a speech community, each with a clearly defined role; one is called High Variety, such as Modern Standard Arabic (HV), and the other is called Low Variety (LV), such as Tunisian Arabic.). The Arabization programs that swept North Africa bolster Tunisia’s constitution laws, namely, Article 1, which dictates that

“Tunisia is a free, independent, sovereign state; its religion is Islam, its language Arabic, and its system is republican.” (https://www.constituteproject.org/constitution/Tunisia_2014.pdf, 4.) However, if Berber enters the orbit of language planning, the Berber language needs to go through the “unplanned” language policy stage – using Baldauf’s (1993-1994) term. Baldauf believes that “planned,” and “unplanned” language aspects can coexist. Baldauf (1993-1994: 83) says that, “If one of these languages is involved in ‘planning’, then a knowledge of the other language situation will be essential to doing work in the planned language.”

Tunisia’s language policy does not allow Berber to be part of the dominant narrative. Yet, language policy is not solely a top-down approach as individuals and communities can adopt their own policies formulated in their attitudes – positive or negative – that are usually unconscious, and as such covert (Sallabank 2013). Language practices among families and communities may also be influenced by “folk linguistics,” where beliefs about language use can affect how language can be used (Nieltzielski and Preston 2003).

Indeed, the Berber language’s convergence under the umbrella of conserving Tunisia’s cultural heritage is problematic. While the country appreciates the diversity of its fabric, accentuated, for instance, through ecotourism, it does not work in favor of the Berber language and its culture, as it makes them static and only a vestige of a hardly recognizable past. Interestingly, parallels can be drawn between the Berber language and Breton, the language of Brittany in France, in political, socioeconomic, and cultural terms. According to Kuter (1989: 76), the political variable refers to the dichotomy between the national versus the regional. In socioeconomic terms, the French language is the language of civilization, whereas Breton is “the language of the past, fit only for backward peasants” (Kuter 1989: 76). As for the cultural variable, while French is considered to have an urban and international status, the Breton language is treated as local with rural character (Kuter 1989: 76). Synonymously, it is not an envious position for the Berbers to find themselves in. This conundrum would raise the question about whether Berber in Tunisia is considered an endangered language.

Language endangerment forms part of the narratives of what is often termed in sociolinguistics circles as “language ecology.” It is defined as “a dynamic network of relationships and interdependencies between the sociocultural, economic, and environmental contingencies that impact on the use, function, structure, and ways of the meaning of languages and their speakers” (Wendell and Heirich 2012: 147). This encompassing definition clarifies that in failing to consider all the variables mentioned above that affect language, it enters the orbit of language death. Furthermore, while language death may not be sudden, i.e., in such cases as a genocide as regards a whole population, consequently, the death of that undocumented language’s last speaker, the process may take years or centuries through a continuum of language contact, attrition, and language shift. Language ecology can take different forms: “stable” and “competitive” (Wendell and Heirich 2012: 148). As the name suggests, “stable” ecologies refer to a multilingual situation where languages interact in an “equilibrium state” without one trying to overshadow the other (Wendell and Heirich 2012: 148).

In contrast, languages in a competitive ecology refer to languages that compete with each other; usually, one is considered the superstratum while the other is the

substratum but classified in various other types of ecologies, namely, “replacement ecologies,” “exploitation ecologies”; and “glocalisation ecologies” (Wendell and Heirich 2012: 148). The first two types of exploitation and replacement ecologies are relevant to the Berber case in Tunisia. While, in the former, language shift becomes the norm in many Berber speech zones where government language policies favor Arabic at the expense of Berber, in the latter, total replacement of Berber with Arabic is occurring at a very rapid rate, especially among Berber youths and the Berber speakers living in major Tunisian cities, especially with the absence of language transmission among Berber generations.

Indeed, UNESCO (2003) lists “Intergenerational Language Transmission” as one of the six factors in Language vitality assessment (Other mentioned factors include: (2) Absolute number of speaker; (3) Proportion of speakers within the total polution; (4) shifts in domains of Language Use; (5) Response to New Domains and Media; and (6) Availability of Materials of language Education and Literacy” (UNESCO 2003, 7).). For instance, a language is considered endangered when Berber is no longer being learned as the mother tongue by children, especially when children respond in a different language. At the end of the continuum, Tunisian Berber does not fit UNESCO’s “critically endangered” classification because the youngest Berber speakers are *not* in the great-grandparental generation, and Berber is still spoken daily. However, Tunisian Berber oscillates between “unsafe” and “definitely endangered” because the domains of language use among children are limited, and the Berber language is used mainly by the older generation.

All observable signs in various fieldwork studies, including the lack of language documentation, all point to language decay and possibly death. Both, Berber speakers, and the Tunisian government, have a role to play in preserving and revitalizing the languages. The known three reactions to language endangerment are listed by Romaine (2008), which are:

1. Do nothing
2. Document endangered languages.
3. Sustain/revitalize threatened languages

The “do nothing policy” is currently adopted by the Tunisian government because its lens does not recognize its importance for its advancement. Whether one likes it or not, languages can have an economic role, especially when a nation struggles to cope in that facet of its existence. Similarly, on the individual or community level, and as stated earlier, the Berbers living in marginalized areas of the south have deep-seated concerns about unemployment, poverty, and the difficulty of life overall. These are the concerns of most Tunisians regardless of race, age, gender, or ethnicity. This does not contradict their awakened identification with the Berber cause because of the Arab Spring revolution. One wants to emphasize that one cannot talk about freedom in a state of penury or when in state of hunger. It is always easier to point the finger at the Berbers themselves for not taking a stand or failing to be more proactive about reviving their language; however, writing about these issues is much easier from the comfort of an academic gaze.

There is little doubt that Berber language is endangered in Tunisia. The prospect of it dying out in Tunisia is probably inevitable for a combination of reasons. According to Dorian (1981: 69), “the search for a single cause which inevitably leads to language death is futile.” Some of the reasons mentioned by Crystal (2014), which determine language death, may not all apply to the Berber case, such as famine or genocide, but other reasons such as the cultural hegemony of the dominant culture, and globalization, may affect how languages are maintained. Geographical isolation is no longer a safe recourse for the Berbers to preserve their language: e.g., transport to, and from, the main city of Tataouine, for instance, has increased the chances of intermingling with local Arabs, which inadvertently affects the maintenance of Berber.

And the “centralisation of power” – in economic terms – may lead these isolated communities to lose their autonomy (Crystal 2014: 102). Additionally, the infiltration of the television, which broadcasts mainly in Arabic and French, plays a major role in this “cultural homogenisation” (Crystal 2014: 102).

Conclusion

The Berbers in Tunisia have an opportunity, in the wake of Arab Spring, to seek governmental support to recognize their language so that it is mainstreamed in the educational system and, hence, gains true vitality. To achieve this goal, the Berbers themselves, without requiring local or international players, need to argue that preserving their language could only enrich Tunisia’s vivid cultural fabric. The Berber narrative should divorce itself from the divisive rhetoric emanating from both Amazigh associations, and Berber militants with dubious allegiances. As colonial history had demonstrated, colonial powers sought to drive a wedge between what formulates the Arab and Berber identity. Furthermore, the Israeli connection with Berber cause cannot be ignored. For instance, Maddy-Weitzman, a former research fellow at the Israeli Moshe Dayan Center (MDC), advises that Israel should use “non-Arab agents” to advance its interests in North Africa (Maddy-Weitzman 2010: 15). The MDC, on its website, does not belie its ideology of supporting and protecting Israel’s future (For more information on the MDC activities, refer to <https://dayan.org/content/about-moshe-dayan-center-mdc>).

In the same vein, Cheryl Benard (2003)’s RAND report entitled *Civil and Democratic Islam* advocates for the revival of the pre-Islamic civilizations to combat Islamic extremism. It is supported by an ideology that Islam is the root of fundamentalist ideologies and violence. In fact, Islam encourages diversity, as stated in the verse, “O mankind, indeed We have created you from male and female and made you peoples and tribes that you may know one another. Indeed, the most noble of you in the sight of Allah is the most righteous of you” (49:13). With only minor occurrences of the rift between the two ethnic groups, both Arabs and Berbers have coexisted for centuries, and continue to share a uniting Islamic religion with a membership that surpasses ethnic divisions or race.

The Berbers of Tunisia would place themselves at a disadvantage if they involve themselves in contentious politics, as variables such as their low demographic number, and their low socioeconomic status, do not work to their favor. Instead, the Berbers, who have remained in those few speech-zones in southern Tunisia should, through various peaceful channels, unite and work together to initiate ways to maintain their language. Fishman's direct advice – “do not leave your language alone” – can be readapted to suit the Tunisian Berber case, for the Berbers to be proactive, as Berbers alone bear the primary responsibility to pass it to posterity (Fishman 2006). Without maintenance and broader domains of Berber use, the Berber language would descend further down the path of attrition, obsolescence, and toward, possibly, language death.

References

- Afrol News (2023) Berber languages “threatened in Morocco, Algeria.” Afrol News, January 23. <http://www.afrol.com/articles/35843>
- Baldauf RB Jr (1993-1994) “Unplanned” language policy and planning. *Annu Rev Appl Linguist* 14:82–89
- Battenburg J (1999) The gradual death of the Berber language in Tunisia. *Int J Sociol Lang* 13: 147–161
- Belhedi H, Ouessar M (1998) The role of NGOs in promoting drylands ecotourism: an ongoing project in southern Tunisia. *Aridlands Newsletter* 43(Spring/summer) <https://cals.arizona.edu/OALS/ALN/aln43/tunisia.html>. Accessed 15 May 2022
- Benard C (2003) Civil democratic Islam: partners, resources, and strategies. RAND, National Security Research Division, Santa Monica
- Brugnatelli V (1987) Deux notes sur l'état d'annexion en Berbère. In: Jungtraithmayr H, Mueller WW (eds) Proceedings of the 4th international Hamito-Semitic congress. John Benjamins [Current Issues in Linguistic Theory 44], Amsterdam, pp 349–359
- Brugnatelli V (2000) Arabe et Berbère à Jerba. In: Youssi A, Benjelloun F, Dahbi M, Iraqui-Sinaceur Z (eds) Aspects of the dialects of Arabic today. Proceedings of the 4th conference of international Arabic dialectology association (AIDA), Marrakesh, April 1–4. Rabat, AMAPATRIL, pp 169–178
- Crystal D (2014) Language death. Cambridge University Press, Cambridge
- Daoud M (2011) The sociolinguistic situation in Tunisia: language rivalry or accommodation? *Int J Soc Lang* 211:9–33
- Dorian NC (1981) Language death: the life cycle of a Scottish Gaelic dialect. University of Pennsylvania Press, Philadelphia
- Ennaji M (1985) Contrastive syntax: English, Moroccan Arabic and Berber complex sentences. CIP, Königshusen
- Fishman JA (ed) (2001) Can threatened languages be saved? Multilingualism Matters, Clevedon
- Fishman AJ (2006) Do not leave your language alone: the hidden status agendas within corpus planning in language policy. Lawrence Erlbaum Associates, Mahwa
- Gabsi Z (2003) An outline of the Shilha (Berber) vernacular of Douiret (Southern Tunisia). PhD dissertation. University of Western Sydney
- Gabsi Z (2011) Attrition and maintenance of the Berber language in Tunisia. *Int J Sociol Lang* 211: 135–164
- Gabsi Z (2020) “Les dialectes montagnards” [The mountain dialects]: towards redefining the Berber status and language attitudes in post-Arab spring Tunisia. *J North Afr Stud* 27(1):80–103. <https://doi.org/10.1080/13629387.2020.1771311>
- Garrett P, Coupland N, Williams A (2003) Investigating language attitudes: social meaning of dialect, ethnicity and performance. University of Wales Press, Cardiff

- Gerges FA (2015) Contextualising the Arab spring uprisings: different regimes, different revolutions, and different trajectories. In: Gerges FA (ed) *Contentious politics in the Middle East: popular resistance and smarginalised activism beyond the Arab uprisings*. Palgrave Macmillan, New York, pp 1–21
- Giles H (1992) Current and future directions in sociolinguistics: a social psychological contribution. In: McElhinny B, Bolton K, Kwok H (eds) *Sociolinguistics today: international perspectives*. Routledge, London, pp 361–367
- Giles H, Rakić T (2014) Language attitudes: social determinants and consequences of language variation. In: Holtgraves T (ed) *The Oxford handbook of language and social psychology*. Oxford University Press, Oxford, pp 11–26
- Giles H, Bourhis RY, Taylor DM (1977) Towards a theory of language in ethnic group relations. In: Giles H (ed) *Language ethnicity and intergroup relations*. Academic Press, London, pp 307–348
- Jay C (2015) A Berber spring: the breakthrough of Amazigh minorities in the uprisings' aftermath. In: Gerges FA (ed) *Contentious politics in the Middle East: popular resistance and smarginalised activism beyond the Arab springs*. Palgrave Macmillan, New York, pp 331–347
- Kossmann M (2013) *The Arabic influence on northern Berber*. Brill, Leiden/Boston
- Kuter L (1989) Breton vs. French: language and the opposition of political, economic, social, and cultural values. In: Dorian NC (ed) *Investigating obsolescence: studies in language contraction and death*. Cambridge University Press, Cambridge, pp 75–91
- Labiadh S (2017) *Al-'aqliyya al-Barbariyya fi Tūnis: Al-'isti'mālāt al-thaqāfiyya wa al-siyēsiyye lil-majmū'ēt al-Barbariyye fi al-Maghreb al-'Arabī*. Sotomedia, Tunis
- Lloyd M (2005) *Beyond identity politics: feminism, power & politics*. Sage Publications Ltd, London
- Louis A (1975) *Tunisie du sud, ksars et villages de crêtes*. Centre National de la Recherche Scientifique, Paris
- Maamouri M (1983) The linguistic situation in independent Tunisia. In: Payne RM (ed) *Language in Tunisia*. The Bourguiba Institute of Modern Languages, Tunis, pp 11–21
- Maddy-Weitzman B (2010) The limits and potentials of Israeli-Maghreb relations. *IPRIS Maghreb Rev* 4:15
- McKenzie R (2010) *The social psychology of English as a global language: attitudes, awareness, and identity in the Japanese context*. Springer, London
- Motylinski, Adolphe de Calassanti (1897) Dialogue et textes en berbère de Djerba. *J Asiat* 10(151): 377–401
- Nieldzielski NA, Preston DR (2003) *Folk linguistics*. Mouton, Berlin
- Ounali MH (1983) La langue des étudiants. In: Payne RM (ed) *Language in Tunisia*. The Bourguiba Institute of Modern Languages, Tunis, pp 103–138
- Romaine S (2008) Linguistic diversity, sustainability, and the future of the past. In: King K, Schilling-Estes N, Fogle LW, Lou JJ (eds) *Sustaining linguistic diversity: endangered and minority languages and language varieties*. Georgetown University Press, Washington, DC, pp 7–22
- Sallabank J (2013) *Attitudes to endangered languages: identities and policies*. Cambridge University Press, Cambridge
- Sayahi L (2014) *Diglossia and language contact: language variation and change in North Africa*. Cambridge University Press, Cambridge
- Silverstein PA (2012) A new Morocco? Amazigh activism, political pluralism and anti–anti-semitism. *Brown Journal of World Affairs* 18(2):129–140
- UNESCO (2003, March 10–12) *Language vitality and endangerment*. International expert meeting on UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000183699>
- University World News (2019) African Union sets up commission for Tamazught language. University World News, November 14. <https://www.universityworldnews.com/post.php?story=2019111081745330>
- Wendell J, Heirich P (2012) A framework for language endangerment dynamics: the effects of contact and social change on language ecologies and language diversity. *Int J Sociol Lang* 218: 145–166. <https://doi.org/10.1515/ijsl-2012-0062>

Section III

Historical Linguistics and Typology

L'archaïsme linguistique en berbère (kabyle): étude sur un corpus parémiologique

13

Linguistic Archaism in Berber (Kabyle): Study on a Paremiological Corpus

Abdelaziz Berkai

*Win yebyan, ittnadi amek;
win yugin, yeqqar : ulamek
(Hamadache 2004: 35)*

Contents

Introduction: Qu'est-ce qu'un archaïsme?	274
Le corpus	277
L'archaïsme linguistique dans les proverbes kabyles	278
L'archaïsme phonétique	278
L'archaïsme morphosyntaxique	279
L'archaïsme lexico-sémantique	284
Conclusion	287
Annexe	288
Bibliographie	289

Résumé

Nous nous proposons ici d'examiner le phénomène de l'archaïsme linguistique à tous les niveaux de l'analyse linguistique : phonétique, morphologique, syntaxique et lexico-sémantique. Notre corpus est constitué de tous les ouvrages et autres matériaux parémiologiques disponibles, notamment du *Dictionnaire de proverbes kabyles* de Ramdane At Menşur (2010) et *Proverbes et dictons kabyles* de Youcef Nacib (Maison des livres 2002). Nous utiliserons aussi, comme un des

Proverbe kabyle que nous traduirons ainsi : « Qui veut, dit : comment ? [litt. cherche comment] ; qui ne veut pas, dit : c'est pas possible »

A. Berkai (✉)

Faculté des lettres et des langues, Département de langue et culture amazighes, Université de Béjaïa, Béjaïa, Algeria

e-mail: abdelaziz.berkai@univ-bejaia.dz

outils de validation de nos hypothèses, les dictionnaires kabyles disponibles (Dallet, Haddadou, Bouamara, Meniche. . .). Le choix de ce type de corpus est motivé par le fait qu'un bon nombre de parémies relevant d'un état de langue qu'on peut qualifier d'archaïque soient conservées en l'état dans la langue actuelle. D'où l'intérêt d'un matériau où diachronie et synchronie se côtoient dans un même état de langue.

Nous essayerons d'abord de définir précisément le concept d'archaïsme et ses manifestations au triple plan diachronique, diastratique et diatopique, ainsi que les différents « degrés d'archaïsation » (Berkai 2014a, 2015), avec des exemples concrets, et terminerons par l'analyse de notre matériau et les conclusions qui en découlent.

Abstract

We propose here to examine the phenomenon of linguistic archaism at all levels of linguistic analysis : phonetic, morphological, syntactic and lexico-semantic. Our corpus is made up of all the books and other paremiological materials available, notably from Ramdane At Menşur's *Dictionnaire de proverbes kabyles* (2010) and Youcef Nacib's *Proverbes et dictons kabyles* (Maison des livres 2002). We will also use, as one of the validation tools of our hypotheses, available Kabyle dictionaries (Dallet, Haddadou, Bouamara, Meniche. . .). The choice of this type of corpus is motivated by the fact that a large number of paremias belonging to a state of language that can be described as archaic are preserved as they are in the current language. Hence the interest of a material where diachrony and synchrony coexist in the same state of language.

We will first try to define precisely the concept of archaism and its manifestations in the three diachronic, diastratic and diatopic planes, as well as the different “degrees of archaization” (Berkai 2014a, 2015), with concrete examples, and finish by analyzing our material and the resulting conclusions.

Mots-clés

archaïsme linguistique · proverbe · kabyle

Keywords

Linguistic archaism · Proverb · Kabyle

Introduction: Qu'est-ce qu'un archaïsme?

L'archaïsme du bas latin *archaismos* emprunt au grec *arkhaismos*, de *archaios* « ancien », signifie tout ce qui est sorti de l'usage ou en voie de l'être à un moment donné d'une expérience humaine. On peut définir ainsi un archaïsme architectural, par exemple, comme un mode ou un élément de construction qui n'est plus en

vigueur, ou tendant à ne plus l'être, chez une communauté humaine particulière à un moment donné de son histoire. *Taxxamt n tmes*¹ « litt. la pièce/chambre à feu » comportant un foyer *lkanun* creusé en son sol et qui était une permanence dans l'habitat traditionnel kabyle n'est plus dans les maisons d'aujourd'hui qu'un lointain souvenir. *Aserwal n ddu* « pantalon un peu court comportant une sorte de grosse poche pendante entre les deux jambes » que portaient jadis la gente masculine et jusqu'à récemment les vieux et sans doute même aujourd'hui, mais de façon extrêmement marginale, est un archaïsme vestimentaire en Kabylie.

Linguistiquement, l'archaïsme désigne pour Zumthor « deux variétés de faits :

1. résurgence, dans l'usage, d'un signe vieilli, voire disparu depuis quelque temps ;
2. survivance, dans un système, d'un signe ayant appartenu à un système désuet » (1967: 19).

Dans cette définition de l'archaïsme Zumthor distingue entre « résurgence » d'un signe, vieilli ou complètement disparu, et « survivance » d'un signe ayant appartenu à un système désuet. Sablayrolles (2006: 6) propose de dénommer ces deux réalités différemment en gardant « archaïsme pour les mots « anciens, vieillissants » et de nommer *paléologisme* un mot réintroduit après une longue absence et dépourvu de signifié préconstruit et de connotation ». Le paléologisme ne se distinguerait ainsi du néologisme que par le seul fait d'avoir déjà eu une existence dans le passé. Ces auteurs parlent ici surtout de l'archaïsme lexical, *signe* pour Zumthor et plus clairement *mot* pour Sablayrolles, alors que ce phénomène concerne tous les niveaux de la langue.

L'archaïsme non lexical, comme le lexical, se manifeste au triple plan diastatique, diatopique et diachronique. Au niveau diastatique, nous verrons ci-dessous (§ 2.2.3) l'exemple de l'usage de l'aoriste sans préverbe *ad* dans sa valeur aspectuelle de non-réel qui ne subsisterait quasiment aujourd'hui que chez les locuteurs du troisième âge. Au niveau diatopique, l'usage de la modalité d'orientation spatiale *-n* très vivant dans les parlers kabyles occidentaux (Tizi-Ouzou, Bouira...) a complètement disparu aujourd'hui des parlers orientaux (Béjaïa, Aokas...) (Berkai 2015: 140). L'état du nom qui se manifesterait par l'absence de la voyelle initiale relèverait de la diachronie (v. § 2.2.2).

Nous avons introduit la notion de degré d'archaïsation (Berkai 2014a, 2015) en distinguant, en synchronie, entre trois types d'archaïsmes en reprenant en partie Zumthor et Sablayrolles :

1. L'archaïsme « fossile » dont il ne reste que le signifiant. Le signifié n'étant pas conservé. On trouve ce type d'archaïsme, en particulier pour les langues à tradition orale comme le berbère, dans des formes de discours relativement datées (proverbes, devinettes, poésie ancienne, toponymie...). Le mot *meqqa* attesté

¹Nous utilisons ici la transcription usuelle du kabyle en caractères latin (voir en annexe l'alphabet utilisé).

- dans le proverbe *aεeqqa, aεeqqa, ad immed meqqa* « Graine après graine, les gouttes vont s'accumuler » (At Menşur 2010: 18), serait un archaïsme fossile, même si son sens est déductible de celui du proverbe. Il n'est attesté dans aucun dictionnaire kabyle disponible ni dans aucun autre dictionnaire amazigh dont nous disposons. Mais il serait peut-être une adaptation, pour rimer avec *aεeqqa* « graine », du mot *timeqqit* « goutte (liquide). Petite quantité » (Dallet 1982: 509) ;
2. Le paléologisme qui est une sorte d'archaïsme néologique. C'est un type qui concerne surtout les langues à tradition écrite où les signifiés d'archaïsmes ayant disparu depuis longtemps de l'usage sont toujours conservés. Il fonctionnerait à la réception comme un néologisme. Son usage relèverait surtout du style. D'où le terme de « stylème » utilisé par J.-M. Klinkenberg pour dénommer ce phénomène linguistique (1970: 17). Le mot *aqlalas* du titre de la célèbre chanson *ay aqlalas* [æqlælæs] de Mouloud Assam, enregistrée en 1978 à la chaîne 2, reprise et popularisée par Ali Amrane, relèverait en kabyle de ce type d'archaïsme (cf. Berkai 2015, voir aussi ci-dessous § 2.3.1) ;
 3. Enfin, des « survivances » ou des archaïsmes que nous avons catégorisés selon leur ignorance par les locuteurs du 1^{er}, 2^e ou 3^e âge comme des archaïsmes respectivement de 1^{er}, 2^e ou 3^e degré. Le mot *tanalt/tanilt* [θænælt/θænilt] « goûter » serait un archaïsme de 1^{er} degré à Béjaïa et à Raffour (Bouira). Les locuteurs du 1^{er} âge dans ces localités, et plus généralement en Kabylie, utiliseraient l'emprunt au français *lguti* [əljuti]. L'usage de l'aoriste nu dans sa valeur aspectuelle de « non-réel » serait en kabyle un archaïsme de deuxième degré. Seuls des locuteurs du 3^e âge l'utiliseraient encore dans la communication quotidienne. Les autres utiliseraient plutôt *ad + aoriste* à sa place : *win yeeyan ad iruḥ* [win jəʃjæn æð ir^soḥ], au lieu de *win yeeyan (eyu + prétérit) iruḥ* (*ruḥ + aoriste nu*) « celui qui est fatigué qu'il parte »². Et enfin, le mot *tibb^wa* [θibb^wa] [*<tiwwa* « dos »] dont le sens est noté par Mammeri dans *Poèmes kabyles anciens* comme « non connu » (1988: 154) serait un archaïsme de 3^e degré en Kabylie occidentale et orientale (Grande Kabylie et Soummam). Le mot *tiwwa* « dos » est d'un usage courant en Kabylie extrême orientale (Berkai 2014b: 829). Le mot *aguliz* [æjuliz] « reste, résidu » dont le sens est aussi noté par Mammeri dans le même ouvrage comme « non connu » (page 150, note 111) serait aussi un

²On retrouve cette substitution de l'aoriste par “ad + aoriste” y compris dans le discours proverbial :

- *win iqerben ajlal ad idel* « celui qui est près de la couverture se couvre » (Hamadache 2004: 24).
celui être près de (+ participe) couverture couvrir (+ ad + 3^{ème} pers. sing. masc. + aoriste)
- *win ur nēdda si tebburt ad iēddi si tzuliyt* (Ben Taleb 2014: 255) (traduction personnelle : « qui n'est pas passé par la porte (qu'il) passe par l'égout »)
celui passer (+ négation + 3^{ème} pers. sing. masc. + prétérit/accompli) par porte (+ état d'annexion) passer (+ ad + 3^{ème} pers. sing. masc. + aoriste) par égout (voir d'autres exemples ci-dessous § 2.2.3).

Il s'agirait peut-être de proverbes récents ou plus probablement adaptés à l'usage récent de la langue.

archaïsme de 3^e degré dans les parlers de la Kabylie occidentale et orientale. Ce mot qui n'est pas archaïque dans certains parlers de la Kabylie extrême orientale (Tizi n Berber, Ait Smail...) (Berkai 2014b: 361), s'est engagé dans l'archaïsation dans d'autres parlers de la même région. A Kherrata, aux frontières avec la wilaya de Sétif (à 60 km au sud-est de Béjaïa), ce mot serait complètement inconnu des locuteurs du 1^e âge. Il serait au moins un archaïsme de 1^e degré dans le parler de cette localité³.

La pertinence de cette catégorisation de l'archaïsme en degrés tient au fait que la langue berbère, en particulier kabyle, se modifie tellement vite, au contact d'autres « grandes » langues de communication, que les locuteurs des différents âges ne parlent quasiment plus la même langue. La modification ne concerne pas seulement le lexique, où les changements sont massifs, mais même les autres niveaux de la langue.

Voici le schéma de notre définition de l'archaïsme :



Le corpus

Notre corpus est constitué d'ouvrages parémiologiques traitant des proverbes kabyles et principalement :

- *Amawal n yinzan n teqbaylit. Dictionnaire de proverbes kabyles* de Remdan At Menşur, Editions Achab, Tizi-Ouzou, 2010. C'est un ouvrage de 588 pages ayant exploité l'essentiel des travaux sur les proverbes parus antérieurement : Y. Nacib, T. Hamadache, S. Ait Ahmed Slimani, A. Abdesselam. . .
- *Proverbes et dictons kabyles* de Youcef Nacib, Maison des livres, Alger, 2002. Ils en existent au moins deux autres éditions antérieures : Andalouses, Alger, 1990 (citée par Remdan At Menşur) et Editions NECIB, 2012.

³ C'est l'un des résultats d'un master soutenu sous notre direction et traitant des archaïsmes lexicaux dans le parler de Kherrata (v. bibliographie).

Par ailleurs, nous avons utilisé tous les dictionnaires kabyles disponibles comme outils de validation de nos hypothèses : Huyghe (1902–1903), Dallet (1982), Ben Taleb (2012), Meniche (2013), Haddadou (2014), Berkai (2014b), Bouamara (2017)...

L'archaïsme linguistique dans les proverbes kabyles

L'archaïsme phonétique

Cela concerne la réalisation, conservée dans des proverbes et qui serait étymologique, de certains sons ou segments phoniques ayant dans l'usage actuel de la langue des réalisations différentes.

2.1.1. *A nnegr-ik ay ul, tarigt (tarik) tɣal i weyyul* « malheur, ô mon cœur, la selle de cheval a été posée sur l'âne » (intronisation abusive, artifice, « à vieille mule, frein doré ») (At Menşur 2010: 544).

La forme *tarik* [θæriçθ] « selle de cheval », seule attestée aujourd'hui en kabyle⁴, et que l'auteur a ajouté entre parenthèse à la notation de la parémie viendrait donc par assimilation régressive de la palatale sonore [g] par la dentale sourde [t] qui devient à son tour sourde [k]. C'est une assimilation régressive par assourdissement qui n'est pas rare dans ce contexte phonique où une sonore s'assourdit pour anticiper l'assourdissement de la dentale –t :

- la consonne -y suffixée, indice de la première personne du singulier en kabyle (et dans d'autres parlers berbères), est généralement assourdie en –x pour anticiper l'assourdissement du pronom affixe régime direct masculin/féminin –t/-tt : *lsiy* « j'ai porté/mis (un vêtement) », mais *lsix-t/-tt* « je l'ai porté(e) ».
- *uɛwij* [uɛwiʒ] « tordu », mais *tuɛwict* [θuɛwiçθ] « tordue », etc.

Même évolution en chleuh : *tarik* « selle pour mule » (Destaing 1938: 258) ; en tamazight du Maroc central : *tarik/tarišt* « selle de cheval » (Taifi 1991: 579) et dans le parler libyen de Ghadamès : *turikt* « selle de méhari » (Lanfry 2011: 312). En rifain et chaoui, l'assourdissement de la palatale sonore *g* s'est accompagné par sa chuintisation, phénomène bien connu dans les parlers zénètes : *trict* (la forme *trikt* y est aussi attestée) (Serhoual 2002: 486) ; *trikt*, pl. *tricin* (Huyghe 1906: 639). Notons ici cette forme nominale sans le nominalisateur vocalique *a-* (v. ci-dessous § 2.2.2).

En touareg, le même mot *tarigt* a subi un autre type d'assimilation qu'on appelle assimilation réciproque ou amalgame en donnant *tarik* [tarik] (De Foucauld 1951: 1623). C'est aussi une évolution qui n'est pas rare en touareg. *Amzag*

⁴Dans le dictionnaire kabyle-français du Père G. Huyghe (1904: 309), c'est déjà la forme *tarik* qui est attestée.

« sourd... » devient au féminin *tamzak* (< *tamzagt*) (De Foucauld 1951: 1273), *amahey* devient *tamaheq* (< *tamahey*)...

2.1.2. *Lmumnin ad akw hlun, leeṣṣat ad msegln* « Les croyants seront tous rétablis, les mécréants vont s'annihiler (foi et incrédulité) » (At Menṣur 2010: 20).

Dans ce proverbe, le préverbe *ad* qui exprime avec le verbe à l'aoriste qui le suit (*hlu* « guérir ») la valeur de non-effectif/non-réel, est prononcé [æd]. La dentale sonore *d* y est prononcée comme une occlusive, alors qu'elle est articulée aujourd'hui comme une fricative [ð] en kabyle. Cette prononciation est attestée aussi dans une autre expression, elle-même archaïque : *ad akk frunt* prononcée [æd æk^w əfrunt]. Cette expression est utilisée pour dire à quelqu'un qu'on trouve pensif, songeur que tous ses soucis et problèmes se dénoueront bien un jour : « ils se dissiperont tous un jour » (Berkaï 2015: 139).

2.1.3. *Bu meyya yeqqar : awah a mitayen* « celui qui a cent répète : ah, si j'avais deux cent » (avidité, "l'appétit vient en mangeant") (At Menṣur 2010: 363).

La prononciation *mitayen* [mitæjən] (deux cent) dans ce proverbe est archaïque. Ce mot est aujourd'hui prononcé *mitin* [mitin]. Ce serait au moins un archaïsme de 2^e degré, puisque cette prononciation est attestée dans le Dallet en seconde position après celle d'aujourd'hui. Les matériaux de ce dictionnaire sont recueillis depuis les années quarante jusqu'au début des années soixante dix du siècle dernier. Cette prononciation est, par contre, celle qui avait cours au début du 20^e siècle. Huyghe ne note pas dans son dictionnaire celle d'aujourd'hui (*mitin*) : *mitain* et *mitein* (1902-1903: 227). La seconde notation de l'auteur semble cependant se rapprocher de celle d'aujourd'hui. Il faut noter, par ailleurs, que parmi les noms qui notent le duel, empruntés à l'arabe où ce paradigme existe, c'est le seul où le segment phonique *-aye-* se vocalise en *i* : *ɛamayen* « deux ans », *yumayen* « deux jours », *cehrayen* « deux mois »... alors qu'en arabe dialectal ce segment semble s'être complètement vocalisé en *i* : *mitin*, *ɛ'amin*, *y'umin*, *cehrin*...

L'archaïsme morphosyntaxique

Cela concerne certains usages morphosyntaxiques qui n'ont quasiment plus cours dans la langue kabyle d'aujourd'hui. Nous aborderons, comme pour le niveau phonétique, trois exemples. Il s'agira de la « copule » *eg* « faire » ; du signifiant nominal sans nominalisateur vocalique, ainsi que de l'aoriste sans la modalité préverbale *ad*.

2.2.1. *ur tegg tazart, a k-ččen, ur tegg ilili, a k-ğğen* « Ne sois pas figue, sinon tu seras mangé, ne sois pas laurier-rose, tu seras abandonné (vertu du juste milieu) » (At Menṣur 2010: 213).

Cette construction syntaxique rappelle l'énoncé nominal à copule *eg* « faire ; être » attesté dans certains parlers du Maroc central (tamazight) et qui est généralisé et

courant en chleuh : *eg* + nom à l'état libre = *d* (copule d'existence) + nom à l'état libre (en kabyle) : *iga argaz* « c'est un homme » (chleuh) = *d argaz* (en kabyle). Chaker a déjà écrit, mais sans donner d'exemples, que ce type de syntagme prédicatif « est connu à l'état de traces en kabyle et touareg » (1994: 2). Dans les parlers chleuhs aussi, la copule d'existence ou particule de prédication *d* est attestée, mais à l'état de traces « dans plusieurs contextes, et particulièrement dans certaines constructions figées :

- négation après *ur* : *urd* (< *ur* + *d*) *asafar-ad* « ce n'est pas ce médicament » ;
- interrogation après *is* : *izd* (*is* + *d*) *asafar-ad* ? « est-ce que c'est ce médicament ? » ;
- comparaison après *zun* : *zund* (*zun* + *d*) *asafar-ad* « comme ce médicament » (El Mountassir 2019: 73).

L'usage copulatif du verbe *eg* est toujours attesté en Kabylie occidentale (Tizi Ouzou, Bouira...), mais quasiment dans les seuls énoncés interrogatifs du type : *amek iga?*⁵ « il est comment, comment est-il fait ? à quoi ressemble-t-il ? », *amek iga : d aberkan ney d amellal* ? « litt. comment il est : (il est) noir ou (il est) blanc ? ». La réponse est : *d aberkan* « il est noir » ou *d amellal* « il est blanc ». L'usage de la copule *eg* en est exclu. S'agirait-il des « traces » de l'usage copulatif du verbe *eg* en kabyle évoquées par Chaker ?

En kabyle, c'est la copule verbale *eg* qui est archaïque, en chleuh c'est plutôt la particule prédicative *d* qui l'est. Alors, lequel des deux usages serait plus ancien que l'autre en berbère ? A. Akouaou qui affirme que cet emploi de la copule *eg* n'est pas attesté en rifain, relève cependant chez les Ayt Ouriaghel un énoncé où les deux copules (*d* et *eg*) coexistent :

- « *ur ggiy d imejni* « je ne suis pas c'est une étoile », « je ne suis pas une étoile » (1979: 113).

Dans cet énoncé, c'est la copule *d* qui est clairement redondante. Ce qui montrerait que dans un état antérieur de ce parler, ce serait probablement la construction sans cette copule qui serait en usage : *ur ggiy imejni*. Un énoncé « témoin » qui montrerait peut-être le passage de l'une vers l'autre dans le rifain et probablement même ailleurs, d'autant que le touareg, considéré comme l'un des parlers les plus conservateurs du domaine berbère, « n'emploie pas la particule prédicative [d], qu'il ignore presque totalement » (Galand 2013: 318).

2.2.2. *Yir wal iqebber am dihan* « Un outrage étouffe comme la rate (effet désastreux des injures et humiliations) » (At Menşur 2010: 566). *Yir wal*

⁵En Kabylie orientale, dans la région de Béjaia en particulier, on utilise le même verbe dans ce contexte, mais à la forme passive : *amek yemmugg/yemmug* (*g* étant articulée comme une occlusive). *Yemmug yakan* « il est déjà fait », *yemmug d amellal* « litt. il est fait il est blanc »/« il est de couleur blanche »...

yessenday urfan « une parole fâcheuse baratte les courroux (une agression, une offense blessent durablement). *Yir wal ma teğğid-t, renyid-t, ma terriđ-t, teḥyid-t* « tu tues une parole détestable si tu l'ignores, tu la maintiens en vie si tu réagis (« seul le silence est grand ») (At Menşur 2010: 567). Beaucoup d'autres mots ne portant pas la voyelle initiale sont attestés dans d'autres proverbes : *yir mgud* (< *amgud*), *mdan* (< *amdan*), *serdun* (< *aserdun*), *rfiq* (< *arfiq*)...

La forme *wal* (< *awal* « mot ») sans la voyelle initiale serait la forme ancienne de l'état d'annexion en berbère selon Chaker. L'auteur donne des exemples de mots composés anciens où le second composant nominal, en position de déterminant, ne porte aucune marque (voyelle) initiale : *iyesdis* « côte » (< *iyes-(i)dis* « os-côte », *asyersif* « aulne » (< *asyer-(a)sif* « bois-rivière » (Chaker 1996: 49). Brugnatelli, Prasse et Galand ont fortement réfuté cette hypothèse. Pour Galand « rien ne permet de croire que l'élément **a**, une fois préfixé au nom, lui ait conféré la valeur de « défini » (...). Il serait étrange que, dans une phase ultérieure, l'élément **w** se soit trouvé associé au « non défini » (...) on ne voit pas pourquoi, au départ, le nom serait « toujours « non défini » quand il est déterminant » (Galand 2013: 137). L'hypothèse la plus ancienne et la plus partagée concernant la voyelle initiale, soutenue d'abord par Stumme (1899), reprise par Laoust (1920) et développée et vulgarisée par Vycichl (1957), soutient que la forme sans voyelle initiale serait celle du nom à l'état libre. La voyelle ou la syllabe initiale serait un « démonstratif »⁶.

Dans le proverbe suivant le même mot est attesté avec la voyelle initiale dans le même contexte :

- *Yir ljerḥ iteqqed iħellu, yir awal iqqaz irennu* « Une plaie se cicatrise et guérit, mais une parole acerbe creuse davantage la blessure » (Nacib 2002: 215).

⁶Le figement de ce démonstratif (*a-*) comme nominalisateur aurait entraîné une nouvelle opposition « défini » vs « indéfini » où la définitude serait portée par d'autres déictiques : *w(a)-* pour le masculin et *t(a)-* pour le féminin : *wargaz* « défini » vs *argaz* « indéfini ». C'est l'hypothèse de K. Prasse (Chaker 2018 : 14) qui pourrait expliquer la naissance de l'état d'annexion masculin en *w-* des parlers du nord. Le figement concerne aussi le « déictique » *wa-* dans beaucoup de nominaux, en particulier dans le domaine de la faune et de la flore, mais qui possèdent souvent des allomorphes ou des variantes qui opposent les deux états. En chleuh : *wagrzam* « panthère », pl. *id-wagrzm*, sans opposition d'état, mais aussi *agrzm*, état d'annexion : *ugrzm*, pl. *igrzmn* (Galand 2013 : 140). En kabyle : *wacnaf* « roquette (bot.) », sans opposition d'état, mais aussi *acnaf*, état d'annexion : *wacnaf* (Dallet 1982 : 88) ; *wazi* « renvois, éructations, aigreurs d'estomac » (Dallet 1982: 883), sans opposition d'état, mais aussi *azi*, état d'annexion *wazi* (Bouamara 2017: 646 ; Meniche 2013: 613), etc. En touareg l'état d'annexion s'exprime par la seule modification de la voyelle initiale *a-* qui se réduit en schwa *ə* ou disparaît :

- *asrir* « terrain plat dur et stérile » (état libre) vs *əsrir* (état d'annexion),
- *tasrirt* « diminutif de *asrir* » (état libre) vs *təsrirt* (état d'annexion),
- *isititen* (pl. de *asiti* « fait d'accroître ») (état libre) vs *sititen* (état d'annexion),
- *tisititin* (pl. de *tasitit* « supplément ») (état libre) vs *tsititin* (état d'annexion) (De Foucauld 1951: 1856, 1877).

Dans la langue courante d'aujourd'hui, la particule *yir* « mauvais » semble être suivie plutôt du nom à l'état libre : *yir awal* « mauvaise parole », *yir tamurt* « mauvais pays ». . . l'état d'annexion est utilisé plutôt par des locuteurs du 3^{ème} âge: *yir wawal*, *yir tmurt*. Il serait donc un archaïsme de 2^e degré⁷, alors que le nom sans voyelle initiale serait un archaïsme « fossile », c'est-à-dire complètement disparu de l'usage.

Un autre archaïsme morphosyntaxique dont nous avons relevé beaucoup d'exemples dans notre corpus est l'usage de l'aoriste sans préverbe. Cet archaïsme serait au moins de 1^e degré dans la plupart des parlers kabyles.

2.2.3. *Wi 'byan taħbult m lenwar, iṣubb s azayar ad yeyleb aydi ttemrit ; wi 'byan lħerma ad tagar yali s adrar ad yečč abelluđ bu tcacit* (proverbe repris dans une chanson de Idir) « qui veut de la galette raffinée, qu'il descende à la plaine et y souffrir comme un chien ; qui veut force dignité, qu'il monte à la montagne et y manger du gland à cupule ».

celui vouloir (+ participe) galette à fleurs, descendre (3^{ème} pers. sing. masc. + aoriste) vers plaine dépasser (+ ad + 3^{ème} pers. sing. masc. + aoriste) chien souffrance ; celui vouloir (+ participe) dignité abonder (+ ad + 3^{ème} pers. sing. fém. + aoriste) monter (3^{ème} pers. sing. masc. + aoriste) vers montagne manger (+ ad + 3^{ème} pers. sing. masc. + aoriste) gland à cupule.

Ce thème nu, sans la modalité *ad*, est aujourd'hui rarement utilisé en kabyle⁸. Il est surtout attesté dans les discours élaborés (contes, proverbes, poèmes anciens. . .) ayant une certaine profondeur historique. Les verbes *ṣubb* [s^ʰobb] « descendre » et *ali* [æli] « monter » expriment ici des procès postérieurs par rapport à celui exprimé par le verbe *byu* [vɣu] « vouloir ». Il s'emploie toujours en énoncé contraint, en

⁷ Même dans les proverbes, comme pour l'aoriste sans particule qui serait lui aussi un archaïsme de 2^{ème} degré (v. ci-dessus note 3), les deux formes du nom (à l'état libre et à l'état d'annexion) sont attestées après la particule *yir* :

- *yir tferni* [état libre : *tiferni*] *am tferyi* « un mauvais choix est comme un manque (échec) » (At Menşur 2010: 566) ;
- *argaz n yir tmeṭṭut* [état libre : *tameṭṭut*], *uṯan-t deg wezniq immut* « le mari d'une mauvaise femme a été découvert mort dans la rue (égoïsme, indifférence, abandon) » (At Menşur 2010: 565).

Mais aussi :

- *yir tagmatt* [état d'annexion : *tegmatt*] *am kalitus, mebeid i yettarra tili* « une mauvaise fraternité est telle l'eucalyptus, il jette loin son ombre (rupture de la solidarité familiale) » (At Menşur 2010: 470).

⁸ Dans certains parlers berbères comme celui du Djebel Nefoussa en Libye ou celui de Siwa en Egypte, ce thème nu sans particule semble avoir complètement disparu (Galand 2013: 225). Il ne serait aujourd'hui « très vivant » que dans les parlers marocains « où il appartient à tous les niveaux de la langue et où je l'ai moi-même observé dans la conversation, il est devenu beaucoup plus rare en Kabylie et en pays touareg, où il se maintient surtout dans les textes littéraires traditionnels – ce qui contribue sans doute à lui donner une saveur stylistique particulière » (*id.*).

subordonnée. Il exprime ici un procès non-effectif, non-réel. Cette valeur aspectuelle est exprimée aujourd'hui en kabyle par *ad* + aoriste (Berkai 2018: 23). Nous avons même retrouvé une partie de ce proverbe dans le dictionnaire de Ramdane At Menşur où le verbe *ali* « monter » est utilisé à l'aoriste accompagné du préverbe *ad* :

- *Win ibyan lherma ad tagwar, ad yali s adrar* « Qui désire force considération, qu'il aille en montagne » (vertu des montagnards) (At Menşur 2010: 34).
- celui vouloir (+ participe) dignité abonder (+ *ad* + 3^{ème} pers. sing. fém. + aoriste), monter (+ *ad* + 3^{ème} pers. sing. masc. + aoriste) vers montagne

Un autre proverbe sur le même modèle :

- *Win ibyan a(d) d-iður lemşam, ad yezwir deg at wexxam* « Qui veut se rendre aux lieux saints, commence par s'occuper des siens » (pèlerinage soufi) (At Menşur 2010: 323).
- celui vouloir (+ participe) visiter (+ *ad* + vers ici + 3^{ème} pers. sing. masc. + aoriste), commencer (+ *ad* + 3^{ème} pers. sing. masc. + aoriste) dans “ceux de” maison (+ état d'annexion)

Ce qui montre que cette évolution n'est pas tout à fait récente dans la langue. Le proverbe suivant avec un aoriste sans préverbe est aussi connu chez des locuteurs moins âgés avec le préverbe *ad* :

- *Mmi-s n tayať yečč-it wuccen* « Que le petit de la chèvre soit dévoré par le chacal » (chacun pour soi) (At Menşur 2010: 367) ou *mmi-s n tayať a(d) t-yečč wuccen*⁹.
- fils-son de chèvre manger-le (+ 3^{ème} pers. sing. masc. + aoriste) chacal (+ état d'annexion)

Ce thème exprime aussi un procès « réel », effectif, toujours en subordonnée et dans le discours élaboré (cf. Berkai 2018):

- *Yebb^wed, yať-n tesša u tessin iwayezniwen* (Dallet) (Galand 2002: 264) « il arrive (arriva) et y trouve (trouva) quatre-vingt-dix-neuf ogres ».
- arriver (+ 3^{ème} pers. sing. masc. + prétérit/accompli), trouver-“vers là-bas” (+ 3^{ème} pers. sing. masc. + aoriste) neuf et quatre-vingt-dix ogres

Cette valeur aspectuelle est exprimée aujourd'hui en kabyle par le prétérit (Basset) ou *accompli* (Galand). On dira, par exemple :

- *Yekcem s axxam yufa-t (af* « trouver » + accompli) *yeččur d lyaci* « il entre (entra) à la maison et la trouve (trouva) pleine de gens ».

⁹C'est la forme que nous connaissons nous-même avant de découvrir l'autre.

- entrer (+ 3^{ème} pers. sing. masc. + prétérit/accompli) à maison trouver-le (+ 3^{ème} pers. sing. masc. + prétérit/accompli) être plein (+ 3^{ème} pers. sing. masc. + prétérit/accompli) “c’est/ce sont” gens

L’archaïsme lexico-sémantique

Il y en a tellement que nous n’avons ici ni le temps ni l’espace pour en analyser ne serait-ce qu’une partie significative. Nous nous contenterons, comme dans les deux parties précédentes, de trois exemples.

- 2.3.1. *Wanag iteddu am wulli, iqqar-as nekk d ilelli* « Alors qu’il vit comme un mouton, il se prétend libéré » (aveuglement) (At Menşur : 543).

Le mot *ilelli* [iləlli] « libre », de même que *tilelli* [θiləlli] « liberté », est un néologisme d’emprunt qui vient du touareg (De Foucauld 1951, III : *éléli* « homme libre » : 1066). Il est attesté depuis les années 1940, en particulier le mot *tilelli*, dans les chants patriotiques de militants nationalistes kabyles, notamment Idir Ait Amrane, mais aussi Laimèche Ali (Achab 2013 : 72). Ramdane Achab qui évoque cet usage ancien raconte que « lors du cours de berbère à l’Université d’Alger, M. Mammeri (2013 : 69) a affirmé que la racine du mot *tilelli* a existé en kabyle, donnant comme preuve l’expression féminine suivante : *a wer tellulliḍ* (puisses-tu ne jamais être libre !) » La forme simple du verbe serait donc *lulli* « être libre », plus proche des formes dérivées *ilelli/tilelli* que de celle du touareg *lullet*, comprenant en plus la consonne radicale *t*. Dahbia Abrous aussi nous a affirmé pendant un cours en 1997/98 à l’Université de Béjaia que cette racine existe en kabyle. Si tel est le cas, et il n’y a aucune raison pour que ce ne le soit pas, le mot *ilelli* « libre » serait donc un paléologisme, c’est-à-dire un archaïsme qui revient dans l’usage sous forme de néologisme. Mohand Akli Haddadou le mentionne comme tel, c’est-à-dire comme néologisme, en 2018 dans son *Dictionnaire des mots nouveaux tamaziyt-français-arabe*, Alger, BERTI Editions : 185.

- 2.3.2. *Usu d aqeccuc, aədil d ameccuc* « Le lit c’est du liège, la couverture, des oripeaux (indigence) » (At Menşur 2010 : 546).

Les deux mots *aqeccuc* [æqəʃʃu] (ici « liège ») et *ameccuc* [æməʃʃu] (ici « oripeaux ») sont surtout connus en kabyle dans l’expression figée *qeccuc meccuc* ayant le sens de « tous les membres de la famille, tout le monde ». *Ffyen-d qeccuc meccuc yer ubrid/s abrid* « tout le monde est sorti dans la rue (du plus âgé au moins âgé) ». C’est une expression très peu connue en kabyle jusque-là, mais qui a grandement gagné en popularité depuis le mouvement populaire du 22 février 2019. Les médias kabylophones (radios et télévisions) s’en sont bien chargés en couvrant les grandes marches populaires hebdomadaires. Beaucoup de personnes comprennent maintenant le sens figé de l’expression, mais ne connaissent pas le sens des deux mots isolément. L’expression semble être plus ancienne que le

proverbe, puisque les deux mots y sont dépourvus de la voyelle initiale, formes elles-mêmes archaïques (v. ci-dessus § 2.2.2). Les deux mots sont attestés dans le dictionnaire kabyle-français de Huyghe : *amcuc* « vieille natte tout usée » (1902–1903: 39) ; *aqcuc* « liège, plaque de liège. On dit aussi *aqeccuc* » (1902–1903: 54). La tension sur la première chuintante dans les deux mots, en particulier dans *ameccuc*, semble être une évolution « récente ». Dans le Dallet (1982: 68), les deux mots sont attestés avec tension sur la première chuintante, mais avec une signification quasi identique : *ameccuc* « vieille natte usée, vieux morceau de natte » et *aqeccuc* « morceau de liège (syn. *aywras*)¹⁰ Feuilles sèches » (Dallet 1982: 643). De tous les dictionnaires du kabyle, celui de Abdel Malek Menniche (2013: 411) est le seul à avoir répertorié cette expression, alors que les deux mots n'y sont pas attestés. Il en donne, dans une entrée spéciale, l'explication suivante : « *Mezzi meqq^wer, irk^wel* » que nous pouvons traduire par « petits et grands, tout le monde ». Elle y est illustrée par l'exemple suivant : *Usan-d qeccuc meccuc* que nous traduirons par : « ils sont venus tous, petits et grands ». C'est bel et bien le sens que cette expression a dans la région de l'auteur, qui est aussi notre région d'origine, où elle est bien connue, en l'occurrence dans la daïra (sous-préfecture) de M'Chedallah dans la wilaya de Bouira. Mais, paradoxalement les deux mots n'y seraient pas connus. D'où leur absence dans le dictionnaire de Menniche. Dans le dictionnaire kabyle-kabyle¹¹ de Kamal Bouamara, originaire de Timezrit dans la wilaya de Béjaïa (Petite Kabylie ou Kabylie orientale), seul le mot *aqeccuc* est attesté, avec cependant un sens différent de celui du Dallet et du Huyghe et évidemment de celui qu'il a dans le proverbe : « *lhila, aselsu, tayawsa* » (Bouamara 2017: 415) que nous traduirons par : « ustensile, habit, chose ». L'expression *qeccuc meccuc* serait vulgarisée d'abord par Berbère Télévision, à travers son correspondant de Bouira, en l'occurrence Ali Yeddou, qui l'utilisait quasiment tous les vendredis en couvrant la marche hebdomadaire du mouvement citoyen dans cette wilaya (préfecture), notamment dans les premières semaines, puis se serait propagée par la suite (ou peut-être simultanément) dans les réseaux sociaux et par la suite dans les autres médias d'expression kabyle. L'expression *qeccuc meccuc* serait donc à l'origine l'équivalent du français « avec armes et bagages », c'est-à-dire en emportant tous les biens qu'on possède, même les plus insignifiants (liège et oripeaux).

¹⁰ Qu'on appelle aussi en kabyle : *aqcur* (Berkai 2014b) (à rapprocher de *iqcer* « écorce », *ax^wnac* (Dallet 1982), mais aussi *iqecci*, prononciation actuel de *iqci* « liège » (Huyghe 1902–1903: 497). Ce dernier auteur donne aussi : *tafernant*, *axunac*, *iḍil* comme signifiants de « liège », en plus de *aqcuc* et *aqeccuc* au sens de « plaque de liège » (*id.*). Le segment *-(u)c* serait donc ici un suffixe expressif ayant le sens synecdochique d'« une partie d'un tout » qui n'est pas très éloigné de l'emploi le plus connu, et surtout vivant, de « diminutif » qu'il a en kabyle (cf. Chaker 1983; Haddadou 1985; Berkai 2007).

¹¹ C'est le premier dictionnaire monolingue du kabyle, paru en 2010 et réédité (édition revue et augmentée) en 2017.

2.3.3. *A bab usigna d waḍu mnee-iyi si nyenbalu* « Ô Maître des nuages et du vent, protège-moi d'une mort par balle ! (horreur des vendettas) » (At Menşur 2010: 368).

Le mot *nyenbalu* [nyənbælu] semble être un composé juxtaposé, même s'il contiendrait la préposition *n* « de », propre aux composés synaptiques, reliant deux composants, nécessairement des nominaux : (*e*)*ny* (< *tinyin*, *timenyiwt* « le fait de tuer, meurtre, assassinat ») + *n* « de » + *balu* « balle » (?). Mais la forme *eny* est celle du verbe « tuer » et le signifiant *balu* n'est pas connu en kabyle. Il n'est attesté dans aucun des dictionnaires de cette langue que nous avons consultés, pas plus que le mot *nyenbalu* lui-même. Balle (à canon) se dit en kabyle *tarşaṣt* [θarʰsʰasʰθ]. Alors, d'où vient cette « curiosité » ? Ce serait peut-être à l'origine, de par sa forme, une création poétique. Les poètes disposeraient de ce que l'on appelle la « licence poétique » qui leur permet de transgresser la norme. Ce serait probablement une strophe à deux vers, un distique qui se serait, le contexte aidant, proverbialisé. Le contexte serait l'époque du colonialisme français où les vendettas étaient relativement courantes.¹² Ce mot signifierait même, le cas échéant, « vendetta », évoquée par l'auteur pour expliquer la signification du proverbe. Les cas de vers, de distiques ou même de tercets qui se sont proverbialisés ne sont pas rares en kabyle. On attribue, par exemple, au grand poète Si Mohand Ou Mhend la paternité du proverbe très connu en kabyle: *ad nerrez wala ad neknu* « nous romprons mais nous ne plierons pas » (plutôt chêne que roseau ; bravoure, volonté, renom) (At Menşur 2010: 427). Youcef Nacib évoque cette attribution en donnant le tercet du poème d'où se serait détaché le proverbe (Nacib 2016: 86), mais écrit plus loin qu'« on dit qu'il remonterait à l'époque de Jughurta » (Nacib 2016: 115):

ad nerrez wala ad neknu “plutôt me briser que m'incliner”
axir deewessu “et mieux vaut la malédiction”
anda tqewwiden ccifan “là où les chefs sont des maquereaux”

Le proverbe suivant : *win yettagwaden yiwen, ur yettagwad ula yiwen* « celui qui craint l'Un, ne craint nul autre (foi) » (At Menşur 2010: 568), serait un distique de Cheikh Mohand Ou Lhocine, cité par Mammeri dans l'ouvrage qu'il lui a consacré (1990: 65). *Yir tagmatt am kalitus, ɣas ɣwezzif messus, mebeid i yerra tili* « Les mauvais frères sont comme l'eucalyptus, bien que haut, il est sans intérêt ; il donne son ombre au loin » que cite Dallet comme étant un chant (1982: 261), At Menşur en cite un distique (le 1^e et le 3^e vers) en guise de proverbe, serait un tercet d'un poème de Si Mohand (Mammeri 1969: 438).

Nous avons relevé beaucoup d'autres exemples intéressants, mais nous n'avons pas suffisamment d'espace pour les aborder ici. Des mots qui n'existeraient qu'en

¹² L'auteur lui-même, Remḍan At Menşur, né en 1937, a vécu cette période du colonialisme français en Algérie. C'est un écrivain et poète dans sa langue maternelle après avoir été un éminent professeur de chimie à l'université. Son « vrai » nom ou plus exactement son nom administratif est Ramdane Ouahes sous lequel il publiait ses ouvrages et travaux scientifiques.

parémiologie, puisqu'à l'usage la langue actuelle semble leur avoir préféré des emprunts. Des emprunts qu'on appellerait « de luxe », puisqu'ils n'enrichissent pas la langue, bien au contraire. En voici des exemples : *kaz* « être prudent, faire attention à », *imerni* (*imernan*) « porte-bonheur », *agrawal*¹³ « effervescence ; bouleversement » (> *tagrawla* « révolution ») ; *mergiged* (*ttmergigid*) « faire des va-et-vient, des zigzags, zigzaguer »...

Conclusion

Nous venons de le voir, les proverbes regorgent d'archaïsmes linguistiques. Ils vont de l'archaïsme fossile aux survivances de divers degrés en passant par le paléologisme. Ils ne concernent pas seulement le lexique qui est évidemment le niveau de la langue le plus touché, mais tous niveaux linguistiques : phonétique, morphologique et même syntaxique.

Au niveau lexico-sémantique, nous avons montré que les mots *ilelli/tilelli* « libre/liberté » pourraient être des paléologismes, c'est-à-dire des mots qui auraient déjà existé en kabyle, et non des néologismes d'emprunt (interne). Nous avons aussi glosé sur l'étymologie et la signification de l'expression *qeccuc meccuc*, en vogue en 2019 pendant le mouvement populaire du 22 février (le *hirak*) dans les médias kabyles, qui aurait étymologiquement le sens de l'expression française « avec armes et bagages ». Nous avons enfin abordé dans cette partie un mot que nous qualifierons, avec Dubois et Marcie (1965: 25) d'« exotisme », c'est-à-dire d'un mot qui se situerait sur l'axe « entre la langue et la non-langue ». Ce mot est *nyenbalu* « la mort par balle » qui aurait plus précisément le sens de « vendetta ».

Au niveau phonétique, nous avons montré que le mot *tarikt* « selle de cheval », attesté avec la palatale sourde *k* dans plusieurs parlers berbères (kabyle, chleuh, tamazight, rifain, chaoui, ghadamsi, touareg...), serait étymologiquement une évolution par assimilation régressive partielle, en l'occurrence un assourdissement de la sonore palatale *g* par la dentale sourde *t* : *tarigt* > *tarikt*. Nous avons, dans un autre exemple, montré la « persistance » d'une prononciation occlusive, qui serait étymologique, de la dentale sonore *d* du préverbe *ad*. Cette consonne est articulée aujourd'hui partout en kabyle comme une fricative [ð]. Par ailleurs, des nominaux empruntés à l'arabe et exprimant le duel, *mitayen* « deux cent » serait le seul en kabyle où le segment *-aye-* s'est complètement vocalisé en *i* : *mitin*. La forme

¹³ *Tagmatt d awal, tasa d agrawal* « La fraternité n'est qu'un mot, l'amour maternel c'est un bouleversement (démensure de la tendresse maternelle) » (At Menšur 2010: 51). La racine *grwl* est donc bien attestée en kabyle et son sens est même plus proche de celui du néologisme *tagrawla* « révolution » (forgé par l'*Amawal* « Lexique de berbère moderne » fait autour de M. Mammeri) et dont on attribue la racine au touareg. R. Achab qui note d'abord l'absence de la racine *grwl* dans les dictionnaires kabyles, relève cependant son attestation dans les expressions : *ay agriwel-iw, ay agriwel-inu a yemma* ! « tombées aujourd'hui en désuétude et remplacées par : *ay axeššar-iw* ! « ô mon malheur ! » (Achab 2013: 145). Cette racine est aussi attestée en Kabylie extrême orientale : « *grawel* (*igrawel, ttegrawal, agrawel*) v. intr. : tomber, dégringoler » (Berkai 2014b: 370).

mitayen, attestée comme variante de la forme principale *mitin* dans le Dallet (1982) et seule attestée dans Huyghe (1904), serait aujourd'hui au moins un archaïsme de 2^{ème} degré en kabyle.

Au niveau morphosyntaxique enfin, nous avons montré l'existence ancienne en kabyle de l'énoncé nominal à copule *eg* « faire ; être » attesté dans certains parlers du Maroc central et qui est courant en chleuh : *eg* + nom à l'état libre, équivalent en kabyle de : *d* (copule d'existence) + nom à l'état libre. *Iga argaz* « c'est un homme » (chleuh) = *d argaz* (kabyle). L'usage copulatif de ce verbe est toujours attesté en kabyle, mais de façon très marginale et dans des constructions très particulières de type : *amek iga ?* « comment est-il fait ? à quoi ressemble-t-il ?... ». Par ailleurs, l'analyse de notre corpus montre que la forme du nom sans le nominalisateur *a-* est bien attestée en parémiologie, mais serait aujourd'hui un « archaïsme fossile » en kabyle et plus généralement en berbère. Un autre archaïsme morphosyntaxique dont nous avons relevé beaucoup d'exemples dans les proverbes est l'usage de l'aoriste sans préverbe. Cet usage serait aujourd'hui au moins un archaïsme de 1^{er} degré dans la plupart des parlers kabyles.

Annexe

L'alphabet utilisé pour la transcription du berbère (kabyle).

Lettres de l'alphabet utilisé	Equivalent en Alphabet Phonétique International (A.P.I)
a	æ/a
i	i
u	u
b	v/β/b
c	ʃ
č	tʃ (affriquée alvéolaire sourde)
d	d/ð (dentale sonore occlusive/fricative)
ḍ	ðʕ (dentale fricative sonore pharyngalisée)
f	f
g	j/ɟ (En kabyle ce sont plutôt les variantes fricatives, qui sont en distribution complémentaire avec les occlusives, qui sont utilisées dans la plupart des contextes phoniques.) (palatale sonore fricative/occlusive)
ḡ	dʒ (affriquée alvéolaire sonore)
h	(laryngale fricative sonore)
ḥ	h (pharyngale fricative sourde)
j	ʒ
k	ɟ/c (palatale sourde fricative/occlusive)
l	l
m	m
n	n
q	q (uvulaire occlusive sourde)
ɣ	ɣ (vélaire/uvulaire fricative sonore)

(continued)

Lettres de l'alphabet utilisé	Equivalent en Alphabet Phonétique International (A.P.I)
r	r/ ^ɾ (vibrante simple/pharyngalisée)
s	s
ʃ	s ^ɸ (sifflante sourde pharyngalisée)
t	θ/t (dentale sourde fricative/occlusive)
ɛ	t ^ɸ (dentale occlusive sourde pharyngalisée)
w	w
x	x (vélaire/uvulaire fricative sourde)
y	j
z	z
ʒ	z ^ɸ (sifflante sonore pharyngalisée)
ɛ	ʃ (fricative pharyngale sonore)

Bibliographie

- Achab R (2013) L'aménagement du lexique berbère de 1945 à nos jours. Tizi-Ouzou, Editions Achab
- Akouaou A (1979) Le verbe g berbère : étude diachronique. Bulletin économique et social du Maroc 140:105–116
- Amawal n *tmaziɣt tatrart (Lexique de berbère moderne) Tamaziɣt-Taɣransist/Taɣransist-Tamaziɣt* (1990). Bgayet: Editions de l'Association Culturelle Tamaziɣt AŽAR
- Ameziane A (2007) Aqlalas ou l'œuvre littéraire qui résiste à l'épreuve de l'usure. <http://tamazgha.fr/Aqlalas-ou-l-oeuvre-litteraire-qui.html>
- At Menšur R (2010) Amawal n yinzan n teqbaylit. Dictionnaire de proverbes kabyles. Editions Achab, Tizi Ouzou
- Ben Taleb B (2012) Dictionnaire français-tamaziɣt. Ccix-iw. Mon maître. Editions Ait Mouloud, Tizi Ouzou
- Ben Taleb B (2014) Amawal n yenzan s teqbaylit. Dictionnaire de proverbes en kabyle. Editions el-amel, Tizi Ouzou
- Berkai A (2007) Lexique de la linguistique français-anglais-berbère, précédé d'un essai de typologie des procédés néologiques. L'Harmattan (réédité par Achab, Tizi-Ouzou, 2009), Paris
- Berkai A (2014a) L'archaïsme en tamazight (kabyle): analyse et identification. In: Dourari A (dir.), La néologie, les corpus informatisés et les processus d'élaboration des langues de moindre diffusion, Timsal n Tamazight, n° 4, p 78-91
- Berkai A (2014b) *Essai d'élaboration d'un dictionnaire tasahlit (parler d'Aokas)-français*, thèse de doctorat, 3 volumes. Université de Tizi Ouzou
- Berkai A (2015) L'archaïsme linguistique : un essai d'identification en berbère (kabyle). In: Yacine T (dir.), Langue et politique en Afrique du Nord, Awal, n°43–44, p 133–144
- Berkai A (2018) Les oppositions verbales berbères à la lumière de la psychomécanique du langage. Studii de Stinta si Cultura XIV 3(54):19–28
- Bouamara K (2017) ISSIN wis sin : Aseggawal n teqbaylit s teqbaylit. Dictionnaire kabyle-kabyle. Editions L'Odyssée, Tizi Ouzou
- Brugnatelli V (2006) L'ancien « article » et quelques phénomènes phonétiques en berbère. In: Dymitr I, Voßen R, Stroomer H (eds) *Etudes berbères III, le nom, le pronom et autres articles*, Actes du « 3. Bayreuth-Frankfurter Kolloquium zur Berberologie », 1-3 juillet 2004. Köppe, Köln, pp 55–70
- Chaker S (1983) *Un parler berbère d'Algérie (Kabylie): syntaxe*. Thèse présentée devant l'Université de Paris V le 16 décembre 1978. Publications de l'Université de Provence

- Chaker S (1994) Chleuh. In: Gabriel Camps (dir.), 13 | Chèvre – Columnnien, Aix-en-Provence, Edisud (« Volumes », no 13), 1994 [En ligne], mis en ligne le 01 mars 2012, consulté le 19 avril 2019. <http://journals.openedition.org/encyclopedieberbere/2270>
- Chaker S (1996) Manuel de linguistique berbère II. Syntaxe et diachronie. Alger, ENAG
- Chaker S (2018) Les marques obligatoires du nom et l'état d'annexion en berbère. In: Edzard L, Sartori M, Cassuto P (eds) *Case and Mood Endings in Semitic — Myth or Reality ? Désinences modales et casuelles dans les langues sémitiques – Mythes ou réalités?* Harrassowitz Verlag [Abhandlungen für die Kunde des Morgenlandes 113], Wiesbaden, pp 162–183
- Dallet J-M (1982) Dictionnaire kabyle-français (parler des Ait Menguellat, Algérie). SELAF, Paris
- De Foucauld C (1951) Dictionnaire touareg-français : dialecte de l'Ahaggar. Tome III et Tome IV, Imprimerie nationale de France, Paris
- Destaing E (1938) Vocabulaire français-berbère, tachelhit du Sous. Leroux, Paris
- Dubois J, Marcie P (1965) Terminologie linguistique (suite). Le français moderne XXXIII:21–33
- Dubois J, Giacomo M, Guespin L, Marcellesi C, Marcellesi J-B, Mével J-P (2002) Dictionnaire de linguistique. Larousse, Paris
- El Mountassir, A. (2019). «Le morphème *d* en tachelhit: auxiliaire de prédication ou préposition de conjonction? », dans *L'ambiguïté catégorielle en Amazigh*, Centre de l'Aménagement Linguistique, Rabat, Maroc. hal-02539150
- Galand L (2002) Etudes de linguistique berbère. Peeters, Leuven/Paris
- Galand L (2013) Regards sur le berbère, Tizi-Ouzou, Editions Achab, 2013
- Haddadou MA (1985) Structures lexicales et signification en berbère (kabyle). Thèse de 3^{ème} cycle de linguistique. Université d'Aix-en-Provence
- Haddadou MA (2014) Dictionnaire de Tamaziɣt (parlers de Kabylie). Alger, Editions BERTI
- Haddadou MA (2018) Dictionnaire des mots nouveaux tamaziɣt-français-arabe. BERTI Editions, Alger
- Hamadache T (2004) Proverbes berbères de Kabylie. Editions Talantikit, Béjaia
- Huyghe G (1902–1903) Dictionnaire français-kabyle. Malines, Belgique
- Huyghe G (1904) Dictionnaire kabyle-français. Deuxième Edition, Imprimerie Nationale, Paris
- Huyghe G (1906) Dictionnaire français-chaoui. A. Jourdan, Alger
- Klinkenberg J-M (1970) L'archaïsme et ses fonctions stylistiques. Le français moderne 38:10–34
- Lanfray J (2011) Dictionnaire de berbère libyen (Ghadamès). Editions Achab, Tizi-Ouzou
- Liazidi N, Oucheni H (2019) Tasleɣt tasnalɣamkant n wawal awessar n tmeslayt n Xerrata d userwes akked tmeslayt n Ayt Smacel. Université de Béjaia, Béjaia
- Mammeri M (1969) Les isefra, poèmes de Si Mohand–Ou-Mhand, texte berbère et traduction. François Maspéro, Paris
- Mammeri M (1988) Poèmes kabyles anciens. Laphomic, Alger
- Mammeri M (1990) *Inna-yas Ccix Muhend. (Cheikh Mohand a dit)*. N.P
- Meniche A-M (2013) Asegzawal n tmaziɣt. Dictionnaire amazigh (Tamazight-Français). Editions Asirem, Bouira
- Nacib Y (2002) Proverbes et dictons kabyles. Editions Maison des livres, Alger
- Nacib Y (2016) Aspects du conte et du proverbe amazighs. Editions Zyriab, Alger
- Sablayrolles J-F (2006) « Terminologie de la néologie : lacunes, flottements et trop pleins. Syntaxe et sémantique n° 7 (Franck Neveu éd.), Presses universitaires de Caen, p. 79–89. http://hal.archives-ouvertes.fr/docs/00/15/44/01/PDF/HAL_JFS_20_2006c_.pdf
- Sablayrolles J-F (2010) Archaïsme: un concept mal défini et des utilisations littéraires contrastées. In: Himy-Pieri L, Macé S (eds) *Stylistique de l'archaïsme*, Collection « Poétique et stylistique ». Presses Universitaires de Bordeaux, Bordeaux, pp 43–65
- Serhoual M (2002) *Dictionnaire tarifit-français*. Thèse de doctorat d'Etat ès lettres. Université Abdelmalek Essaâdi
- Taïfi M (1991) Dictionnaire tamazight-français (parlers du Maroc central). L'Harmattan-Awal, Paris
- Zumthor P (1967) Introduction aux problèmes de l'archaïsme. Cahiers de l'Association internationale des études françaises 19:11–26

Reconstructing the Proto-Berber Verb System in Afro-Asiatic Perspective

14

Vit Bubenik

Contents

Introduction	292
Berber Aspectual System	292
The Role of Ablaut in the Formation of the Three Aspectual Stems	293
Stative Conjugation	294
Marking for the Contrast of Imperfectivity Versus Perfectivity	295
Temporal/Aspectual Contrasts	296
Lexical Aspect Contrasts (Initial Phase, Repetition of an Event)	297
Conjoining Clauses in Berber and Semitic	298
Berber Diathetic System	298
Semitic Parallels in Marking for the Imperfective	300
Reconstructing Common Proto-Semito-Berbero-Cushitic	303
Conclusion	304
References	305

Abstract

The central piece of this contribution is the reconstruction of the Proto-Berber aspectual and diathetic system. On the basis of Tuareg, Tamazight, and Kabyle the basic imperfective form of Proto-Berber is reconstructed with the second radical reduplicated **yi-kārras* “he is binding”; two monosyllabic forms with different vocalic pattern are found in the formation of “aorist” (unmarked) and perfect, **ya-kris* and **yu-kras*. The dichotomy of the disyllabic imperfective versus the monosyllabic aorist and perfect can be projected back to a Common Afro-Asiatic stage where several disyllabic imperfective forms (with or without a lengthened radical vowel, and with or without the reduplication of C2) *ya-CaCaC*, *ya-CaCCaC*, *ya-CāCaC*, *ya-CāCCaC* can be reconstructed. As in Semitic and Old Cushitic, the workings of ablaut are observed in both the radical, and thematic vowel (cf. Akkadian *i-parras* “he separates” vs. intensive

V. Bubenik (✉)

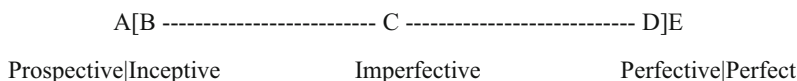
Faculty of Arts and Social Science, Memorial University of Newfoundland, St. John’s, NL, Canada
e-mail: vbubenik@mun.ca

(imperfective) *u-parras* vs. intensive perfective *u-uparris*): Beja (Old Cushitic) *danbiil* “he collects,” *?i-dbil* “he collected,” *?ii-dbil* “he has collected”). During their history, the Afro-Asiatic languages diverged considerably in the renewal of the marked category of perfect.

Introduction

A number of theoretical frameworks are available for the analysis of tense and aspect (available in Comrie 1976; and more recently in Binnick 2012). This research uses the “chronogenetic” approach developed in our systemic analysis of tense and aspect in Indo-European languages (Hewson and Bubenik 1997). We understand aspect as the representation of the time involved in the event time, for which there are five cardinal positions: prospective, inceptive, imperfective, perfective, and perfect (retrospective) as shown in (1):

- (1) Event time with five ‘trappings’



The well-known opposition of perfectivity (completeness) can be marked morphologically in two ways in terms of markedness: perfective (complete) versus imperfective (incomplete) as marked versus unmarked, or unmarked versus marked. The first strategy is exploited in Slavic languages by means of “preverbs” (prepositions) in pairs, such as *on šěl* [šol] “he went” (imperfective), *on pri-šěl* [pri-šol] “he came” (perfective) in Russian. The second strategy is available in Berber (and other Afriasiatic languages) whereby the imperfective member of the opposition is marked iconically by enlarging its stem by various means (see section “[The Role of Ablaut in the Formation of the Three Aspectual Stems](#)”).

Berber Aspectual System

Unlike Semitic languages, personal markers in Berber languages are distributed in one fashion across the whole spectrum of their three aspectual stems: unmarked (aorist), imperfective (incomplete), and perfect (complete). Their distribution is as shown in (2):

- (2) The distribution of personal affixes in Berber (Tamazight)
- | | | |
|--------------|---------------------|---|
| Prefixes: | 3rd sg. and 1st pl. | y- and n- |
| Circumfixes: | 2nd sg. and 2nd pl. | <u>t</u> - <u>d</u> and <u>t</u> - <u>m</u> |
| Suffixes: | 1st sgs and 3rd pl. | -x and -n |

The situation in the plural sub-paradigm is an eloquent example of morphological economy (in the sense of Carstairs-McCarthy 1992), whereby the same nasal affix indicates the 1st pl. as a prefix, but the 3rd pl. as a suffix. Their uniform distribution across the three aspectual stems in Tamazight is exemplified in (3) with French terminology in brackets:

- (3) The distribution of personal affixes in the three aspectual stems in Tamazight
- | | | | |
|-------|--------------------|----------------------------|-----------------------------------|
| | Unmarked | Imperfective | Perfect |
| | ["aoriste"] | ["inaccompli"] | ["accompli"] |
| sg. 1 | af-əx
"I found" | tt-afa-x
"I am finding" | ufi-x
"I found" ~ "have found" |
| 2 | t-af-əd | tt-afa-d | t-ufi-d |
| 3 | y-af | i-tt-afa | y-ufi |

The Role of Ablaut in the Formation of the Three Aspectual Stems

Ablaut plays a considerable role in the formation of the three aspectual stems unmarked (aorist), imperfective (incomplete), and perfect (complete), but it is difficult to ascertain some significant semantic correlates with various lexical categories. As in Semitic, there is a broad dichotomy of the verbs of action versus those of states. The vocalic pattern of the verbs of action usually contrasts the unmarked with the perfect (*y-af* "he found" vs. *y-ufi* "he has found"), while the vocalic pattern of stative verbs does not change (*y-užəd* is both the unmarked and the perfect "is/was/has been tired"). As shown in (4), the vocalic pattern of the negative perfect of stative verbs is identical with that of the perfect (and the same pattern is found with the passive of active verbs):

- (4) Aspectual system of Tamazight of the Ayt Ndhir (Penchoen 1973)
- | | | | | | |
|---------|----------|--------------|---------|------------------|------------|
| | Unmarked | Imperfective | Perfect | Negative Perfect | |
| State | užəd | tt-užəd | užəd | užəd | "be ready" |
| Action | aməz | tt-aməz | uməz | umiž | "take" |
| Passive | tty-imz | tty-amaz | tty-imz | tty-imz | "be taken" |

There are four "patterns" (corresponding to the IE notion of "conjugation" as in Latin) of the formation of the perfect stem:

- Ablaut of initial *a* > *u* (or *a* > *i*)
"take" *aməz* > *uməz*, "carry" *awəy* > *iwəy*
- Ablaut (of initial or internal *-a-*) plus *i/a* alternation
"find" *af* > *ufi*, "be born" *lal* – *luli*
The alternant with *-i* is used with participants in discourse (1st and 2nd Sg.):
ufi-γ "I have found," *t-ufi-t* "you have found" versus *y-ufa* "he has found," and *n-ufa* "we have found"

- Alternation only with verbs in *-u* and *-C*
 “go, walk” *ddu* > *ddi/a*
 “give” *š* > *ši/a*, “be, do, make” *g* > *gi/a*
 “eat” *čč* > *čči/a*, “kill” *nəγ* > *nyi/a*
- Partial reduplication (gemination) only with “qualitative” verbs
 “get big” *mɣur* > *məqqur*, “get small” *mziy* > *məzziy*
 “get white” *mlul* > *məllul*

Stative Conjugation

That (3) does not represent an original state of affairs is clearly indicated by the situation in Tuareg and Kabyle where the stative verbs are inflected only by suffixes, very much as in Ancient Semitic (Akkadian) and Egyptian.

(5) Stative conjugation in Egyptian, Akkadian, and Berber (Kabyle and Tuareg)

	Egyptian “be satisfied”	Akkadian “separate”	Kabyle “be famous”	Tuareg “to be in a small quantity”
sg. 1	ḥatp-áku	pars-āku	mšhur-γ	
2m	ḥatp-áta (~ ḥatáp-ta)	pars- āta	mšhur-ḏ	
f	ḥatp-āti (~ ḥatáp-ti)	pars- āti		
3m	ḥatp-a	paris	mšhur	dərus
f	ḥatp-áta (~ ḥatáp-ta)	pars-at	mšhur-t	

In Kabyle the “quality” verbs (ca. 60 of them are listed in Chaker 1983: 117–118) allow for the contrast stative versus progressive (formed as imperfective) as shown in (6), quoted after Mettouchi (2004: 106):

(6)	mqq ^a r-t be-big-sg.3/Fem. “the girl was big”	təqšišit girl	(Kabyle, Mettouchi 2004)
	di šətwə in winter	tti-mɣur-n PROG-be-big-Pl.3/Masc.	wuḏan nights
	‘in winter the nights (are) grow(ing) longer’		

In Central and South Semitic languages, the stative conjugation was recategorized as the perfect conjugated by direct (subject) pronominal suffixes: *-ku sg. 1, *-ta sg. 2/Masc. *-ti sg. 2/Fem. Their original distribution was “obscured” by non-proportional analogy in Gəʕəz (*labas-ku* “I dressed,” *labas-ka* “you (Masc.) dressed” *labas-ki*, “you (Fem.) dressed”), and Arabic *labis-tu*, *labis-ta*, *labis-ti*.

Marking for the Contrast of Imperfectivity Versus Perfectivity

In Berber the imperfective member of the opposition of perfectivity (perfective/complete – imperfective/incomplete) is marked “iconically” – with the imperfective being the marked category – by various means: i) reduplication of the first or second radical, (ii) insertion of *-a-* in its stem, and (iii) prefixation of *tt-*. It is important to acknowledge that the imperfectives are derived from the unmarked category (aorist) as seen in the following examples from Kabyle (Galand 1987):

- (7) Basic aspectual forms of in Kabyle (based on Galand 1987)
- | Unmarked | Imperfective | Perfect | Negative Perfect |
|---------------------------|-----------------|---------------|------------------|
| “plow” (ad) <i>y-krəz</i> | <i>i-kərrəz</i> | <i>y-krəz</i> | <i>y-kriz</i> |
| “come” (ad) <i>yəddu</i> | <i>i-ttəddu</i> | <i>y-ədda</i> | <i>y-əddi</i> |

- (8) Formation of the imperfective stem in Tamazight (after Penchoen 1973: 34–39)
- i-gən* “he slept” versus *i-ggan* (imperfective)
i-krəz “he plowed” versus *i-kərrəz* (imperfective)
i-skr “he did” versus *i-skar* (imperfective)
y-aməz “he took” versus *i-tt-aməz* (imperfective)

The “relative mode” allows only for the binary contrast of the imperfective versus perfect. It is formed by the participial suffix (*-n* in sg. and *-nin* in pl.).

- (9) Relative mode in Tamazight
- | | Imperfective | Perfect | Negative Perfect |
|-----|-------------------|-----------------|------------------|
| sg. | <i>i-kərrəz-n</i> | <i>i-krəz-n</i> | <i>i-kriz-n</i> |
| pl. | <i>kərrəz-nin</i> | <i>krəz-nin</i> | <i>kriz-nin</i> |

These forms are used when the verb is in a subject relative clause: *i-kərrəz-n* “who ploughs/will plough/is ploughing” versus *i-krəz-n* “who (has) ploughed.” As shown in (10) relative clauses are usually introduced by a demonstrative particle (*nna* in Tamazight and Tashelhiyt). There are considerable differences in the agreement pattern of the participle across the dialectal spectrum (Kossmann 2003). The following examples from Tamazight (Penchoen 1973: 70) show agreement of the participle of the copula *g* “be, do” with its antecedent in number.

- (10) *aryaz (nna) iga-n aməqqran* (Tamazight, Penchoen 1973: 70)
 ‘the man who is big (or ‘important’)
irəyzan (nna) ga-nin iməqqran
 ‘the men who are big (or ‘important’)

If the relativized clause expresses possession the participle of the verb of existence *ili* (“be, exist”) is used:

In Tamazight further distinctions are expressed by means of grammaticalized lexical auxiliaries:

- (a) “go” > future time reference
 - (b) “be, exist” > stativity (equational predication)
 - (c) “be,” “exist” > progressive aspect
 - (d) “pass (by)” > habituality
 - (e) “stay, sit” > continuative
 - (f) “stand up” > subitaneity
-
- (a) “Future”: perfect of the verb *ddu* “go” followed by the projective particle with the main verb in the unmarked: *idda ad-iddu* “he is going to go”
 - (b) “Stativity”: perfect of the verb *ili* “be, exist” followed by the perfect of a stative (qualitative) verb: *i-mɣur* “he/it is big” > *i-məqqur* “he/it gets big” > *illa i-məqqur* “he is grown up”
 - (c) “Progressive” aspect: perfect of the verb *ili* “be, exist” in combination with the imperfective stem of an action verb: *illa i-kərrəz* “he was ploughing.” In Tashelhiyt the progressive aspect is indicated by the particle *ar* (*ar i-kkrz* “he is ploughing”)
 - (d) “Habituality” in the past is indicated by the perfect of the verb *kk* “pass (by)” with the perfect of the main verb: *kki-x ssən-x* “I used to know”
 - (e) “Continuative”: perfect of the verb *qqim* “stay, sit” in combination with the imperfective stem preceded by the particle *ar*. Contrast the continuative *i-qqima ar i-ttə-ddu* “he kept on walking” with the progressive introduced by the extensive particle *la i-ttə-ddu* “he is going.”
 - (f) “Subitaneity”: the unmarked of the verb *kkər* “stand up” with the main verb in the unmarked or the perfect: *i-kkər uryaz y-aməz-t* (unmarked) ~ *y-uməz-t* (perfect) “(so) then the man seized him” (The term “subitaneity” or “suddenness” may not be the best term to denote “the initiation of a new action” in the narrative string).

Lexical Aspect Contrasts (Initial Phase, Repetition of an Event)

The notion of “inceptivity” (expressing the initial phase of an action/event) is realized by means of lexical verbs of Arabic origin: *zayd* “start” (< Ar *zād* “grow, increase, multiply”) or *bdda* “start” (< *badā* “become evident, appear”). (CONT “continuative” IPFVE “imperfective” in parsing):

- (13) *zaydn* *ar* *t* *kkatn* (Tashlehiyt, Stroomer 2002: 4)
 start-pl.3 CONT him IPFVE-beat-pl.3
 “They started to beat him”

As in many languages, the passive is normally used without the agentive phrase:

- (19) y-krz yigr (Kabyle, Chaker 1983: 310)
 “he cultivates the field” >
 y-ttw-akrz yigr
 “the field is/was/has been cultivated”
 y-amz izmawn (Tashlehiyt, Stroomer 2002: 212)
 “he caught lions” >
 tty-amzn izmawn
 “the lions are/were/have been caught”

The agentive phrase can be added by means of the instrumental preposition *s* “with.”

In Tamazight, aspectual contrasts with the perfect are limited depending on the verbal pattern in (20).

(20) Aspectual contrast in the passive, in Tamazight (Penchoen 1973)

	Unmarked	Imperfective	Perfect
(a) “take”	aməz	tt-aməz	uməz
Passive	tty-imz	tty-amaz	tty-imz
(b) “ask”	sal	tt-sal	sal
(Passive forms were not provided)			
(c) “kill”	nəγ	nəqqa	nyɪ/a
Passive	ttu-nəγ	ttu-nəγ-a	ttu-nyɪ/a
(d) “plough”	<u>krəz</u>	<u>kərrəz</u>	<u>krəz</u>
Passive	ttu- <u>krəz</u>	ttu- <u>kraz</u>	ttu- <u>krəz</u>

In (a) *tty-imz* indicates both the unmarked “he was taken” and the perfect “he has been taken” while in (c) this contrast is viable (*i-ttu-nəγ* “he was killed” vs. *i-ttu-nyɪa* “he has been killed”). In (d) both active and passive are based on a simple binary opposition of imperfective versus perfect. In the case of “ambitransitive” verbs (such as *krz* “plough”), there is a contrast of “resultativity” whereby the perfect form states a fact while the passive form indicates that this fact is the result of a previous action instigated by an agent. This is shown in (21) by the following minimal pair in Kabyle (quoted after Mettouchi 2004: 102):

- (21) lgər nni y-krəz (Kabyle, Mettouchi 2004)
 field (anaphoric) sg.3/Masc.-plough-PERF
 “the field (in question) is ploughed”
 lgər nni y-twa-krəz
 field (anaphoric) sg.3/Masc.-plough-PASS-PERF
 “the field (in question) has (finally) been ploughed”

An earlier state of affairs is found in Tuareg, which allows for a three-way distinction of aspect in the passive. The ablaut variants *akraz* and *ikraz* “be gained,” in combination with two passivizing prefixes *t-* and *tiw-* (unmarked *t-akraz/tiw-ikraz*, perfect *t-ikraz/tiw-akraz*), versus the imperfective featuring the geminated dental prefix (*tāt-akrāz/tīt-ikrīz*) yield the three-way aspectual contrast (Note: *ā* with macron stands for *â* in Prasse (1973: 86–89)).

Semitic Parallels in Marking for the Imperfective

The affinity of the *inflectional* imperfectivizing prefix *ttw-* and the *derivational* passive marker *ttu-* (in Kabyle) is intriguing. According to Voigt (1987), the imperfectivizing prefix *t-* of Berber has a functional parallel in the Akkadian forms with “iterative-habitative” infix *-tan-*. Its *-t-* functions as the derivational affix in the formation of the reflexive/reciprocal (*pi-t-šuš-um* “to anoint oneself,” *mi-t-gur-um* “to agree (with one another)”), but also as the inflectional aspectual marker in the formation of the perfect *i-p-ta-ras* “he has separated.” In the Proto-Berber forms based on Tuareg (as reconstructed by Prasse 1973: 86–89) the prefix *t-* derives the passive stem **t-akraz* “be gained” (unmarked) and the perfect **t-ikraz* (with ablaut). The imperfective form

**tāt-akrāz* reduplicates the prefix *t-* for its double function of the passive and the imperfectivizing aspectual marker (with lengthening of the radical vowel).

In Akkadian the iterative infix *-tan-* is found across the four aspectual categories (in all their derivatives) and quasinominal forms (participles, infinitives).

(22) Akkadian iterative forms

	Perfective	Imperfective	Perfect	Stative
“separate”	i-prus	i-parras	<u>i-p-ta-ras</u>	paris
Middle	<u>i-p-ta-ras</u>	<u>i-p-ta-rras</u>	i-p-ta-t-ras	pi-t-rus
Iterative	*i-p-tan-ras	i-p-tan-arras	*i-p-ta-tan-ras	pi-tan-rus
	> <u>i-p-ta-rras</u>		>i-p-ta-tar-ras	

In iterative *-n-* assimilates to the second radical, resulting in the homophony of the iterative perfective and the middle imperfective (*i-p-ta-rras*). Another ambiguous form is *i-p-ta-ras* (underscored), which indicates the active perfect (here the allomorph *-ta-* is an inflectional aspectual marker) and middle perfective (here the infix *-t(a)-* is a derivational element forming the middle voice stem). In the middle perfect the infix is “geminated,” **i-p-tat-ras* with *-ta-* (perfect) followed by *-t-* (middle), yielding another ambiguous form *i-p-ta-rras* (see von Soden 1952: 120–121, for documented forms). The doubling of the infix *-t-* in the middle perfect in Akkadian (**i-p-ta-t-ras*) has a parallel in the passive imperfective in Tuareg: *ittā-krāz* “be gained,” imperfective *itātā-krāz* < Proto-Berber **yi-t-ikraz*, imperfective **yi-tāt-akrāz*. (*æ* stands for Prasse’s *ä*)

As mentioned in section “[Marking for the Contrast of Imperfectivity Versus Perfectivity](#)”, in addition to marking imperfective by the prefixation of *tt-*, there are two other “iconic” strategies:

- (i) reduplication of the first or second radical (*i-gən* ‘he slept’ > *i-ggan* ‘he sleeps’ (imperfective), *i-kṛəz* ‘he plowed’ vs. *i-kərrəz* ‘he plows’ (imperfective)
- (ii) reduplication of the first radical and insertion or addition of *-a-* in/to the imperfective stem (*i-γər* ‘he called’ > *i-qqar* ‘he calls’ (imperfective), *i-nəγ* ‘he killed’ > *i-nəqqa* ‘he kills’ (imperfective).

The trichotomy (“three-way” contrast) of the disyllabic imperfective versus the monosyllabic unmarked versus the ablauted monosyllabic perfect in Northern Berber (Kabyle), and the Atlas Berber (Tamazight and Tashelhiyt) was projected back to Proto-Berber by Prasse (1973, 2003):

(23) Proto-Berber aspectual system

	Unmarked	Imperfective	Perfect	
Kabyle	yə-knəf	i-kannəf	yə-knəf	“cook, bake”
	y-akwr	i-tt-akwr	y-ukər	“steal”
	yə-ls	i-lass	y-əlsa	“dress up”
Tashelhiyt	y-akwr	i-tt-akwr	y-ukwr	“steal”
	i-ls	i-lssa	i-lsa	“dress up”
Proto-Berber	*yaknuf/ *yaknif	*yikannaf	*yuknaf	“cook, bake”
	*yahkur/ *yahkir	*yittahkar (for *yihakkar)	*yuhkar	“steal”
	*yalsih	*yilassah	*yulsah	“dress up”

(Note: In Proto-Berber Prasse marks all the short vowels with the diacritic *ă, ĭ, ŭ*)

The vocalic pattern of the Proto-Berber imperfective **yi-kannaf* ‘he cooks’ (and **yi-kārras* ‘he binds,’ with a long radical vowel) corresponds to the Akkadian imperfective *i-parras* ‘he separates’ and Old Ethio-Semitic (Gəʕəz) *yə-kattəb* ‘he writes.’ The vocalic pattern of the unmarked **yaknuf*/**yaknif* corresponds to the Akkadian perfective (preterit) *i-prus* ‘he separated,’ *i-pqid* ‘he commanded; watched,’ *i-kbit* ‘it was heavy.’ The perfect (**yu-knaf*) is formed by ablaut: (Tamazight) *y-aməz* ‘he held’ > *y-uməz* ‘he has/had held’; (Tashelhiyt) *y-all* ‘he pulled,’ *y-ull* ‘he has/had pulled.’ The pattern of the middle *t*-perfect in Akkadian, *i-p-ta-ras*, is ambiguous between the middle perfective *i-l-ta-baš* ‘he dressed up’ and the active perfect (*i-m-ta-ḥaš* ‘he has hit’). The active interpretation is common with only a small group of verbs denoting ‘holding, grasping, seizing, taking’ (*i-š-ta-bat* ‘he has grabbed’ (Bubenik 2017: 49–52)).

However, an earlier state of affairs is seen in the aspectual system reflecting the dichotomy of verbs of action versus those of state. Relics of this earlier system are

found in Tuareg (*sədid* “be thin”) and Kabyle (*dərus* “be in a small quantity”) where the stative verbs are inflected by suffixes, very much as in Ancient Semitic (Akkadian) and Egyptian (see section “[Stative Conjugation](#)”). In Akkadian the distinction between verbs of action and those of state was indicated by ablaut in the thematic vowel of action verbs versus its absence in those of state (contrast *i-parras* “he separates,” *i-prus* “he separated” with *i-qarrib* (Assyrian) “he approaches,” and *i-qrib* “he got closer”). Thus, an earlier stage of Common Semito-Berber (before the rise of the perfect by grammaticalization of the stative) can be portrayed as shown in (24).

(24) Common Semito-Berber Aspectual System with Stative

	Perfective (Past)	Imperfective (Present)	Stative
Action verbs	*ya-CCi/uC	*yi-Ca/āCCaC	
Stative verbs	*ya-qrib	*yi-qarrib (Assyrian)	*CaCi/uC

In Akkadian the vocalic pattern of the stative *CaCi/uC is identical with that of verbal adjectives: *damq-um* “good” (*damiq-um), Fem. *damiq-tum*, *laber-um* “old”; *marš-um* “sick” (< *maruš-um), Fem. *maruš-tum*, *ḥamš-um* “five,” Fem. *ḥamuš-tum* “five-day week.” The same pattern with the lengthened thematic vowel is found in both East (Akkadian) and West Semitic languages in adjectives and in the passive participle: Aramaic *kəṭīb* “written” (< *kaṭīb), Arabic *marīḍ* “sick,” *xamīs* “Thursday” (= fifth), Gəʕəz *katūb* “written,” and Akkadian (poetic) *karūb-um* “blessed.”

As above, the pattern with ablaut in the perfective (i-CaCCaC vs. i-CCuC) is typical of active verbs, and the pattern without ablaut (i-CaCi/uC vs. i-CCi/uC) is typical of the stem of stative verbs. (Note: In grammars of Akkadian “imperfective” is called “present” (von Soden 1952) or “durative” (Huehnergard 1997)).

(25) The main patterns of active and stative verbs in Akkadian

Imperfective	Perfective	Stative/Verbal adjective
i-parras “he separates”	i-prus “he separated”	paris “it is/was separated”
i-kabbit “it is heavy”	i-kbit “he became heavy”	kabit “it is/was heavy”
i-ḥalliq “he flees”	i-ḥliq “he fled”	ḥaliq “he is/was destroyed”
i-balluṭ “he is alive”	i-blut “he recovered”	baliṭ “he is/was alive/healthy”

There are also isolated examples of the stative with the thematic vowel –a-: CaCaC as in *rapaš* “it is/was broad,” *balat* (Old Assyrian) “he is/was alive/healthy,” *ḥalaq* (Old Akkadian/Old Assyrian) “he is/was destroyed.” From the point of view of subsequent development of West Semitic languages here is precisely an earlier state of affairs in the formation of the stem of the “neo”-perfect (innovative perfect of Hebrew, Aramaic, Gəʕəz, Arabic) by stative > perfect shift.

Formation of the stative stem in West Semitic languages is shown in (26)

- (26) Formation of the stative stem in West Semitic languages
 CaCiC: Arabic *kabid-a* “he wore down,” Hebrew *kābēd* (< *kabid-a) “he was/became heavy”
 Arabic *qarib-a* “he came near,” *qarīb* “near”
 CaCuC: Arabic *qarub-a* “he was near,” Hebrew *qārōb* (< *qarub) “near relation” (adjective)

The pattern CaCuC is relatively well preserved in Classical Arabic: *kabur-a* “he was/became great/old,” *hasun-* “he was/became beautiful,” etc. Later developments of lexicalization (*kabat-tum/kabit-tum* “liver” = i.e., “heavy” organ versus lungs “light” organ (Akkadian), *kabid-/kibd-* “liver” (Arabic)), and analogical leveling with action verbs (Hebrew *qārāb* “he approached” and *pāʿal* “he did”) obscured these morpho-phonemic patterns.

The active “neo”-perfect of West Semitic languages (*katab-a* “he wrote”) is based on the pattern CaCaC documented with several stative verbs in Old Akkadian and Assyrian. In Akkadian the paradigmatic perfect (with stative and some active verbs) was formed by the infix *-ta-* inserted after the first radical: C-ta-CaC (as shown above). It should be noted that traces of this pattern survived in Arabic as measure VIII: *fahim-a* “he understood” > *i-f-ta-ham-a* “he (has) understood,” *samiʿ-a* “he heard” > *i-s-ta-maʿ-a* “he listened.”

Reconstructing Common Proto-Semito-Berber-Cushitic

It remains questionable whether the three-way system comprising the disyllabic imperfective, the monosyllabic unmarked (perfective), and the marked category of the perfect can be projected even farther back to the Common Semito-Berber-Cushitic stage (within the Afrasian phylum). Consider the parallelism with the three aspectual categories in Old Cushitic (Beja); data in Voigt (1987) (Note: Appleyard (2012) labels perfect “aorist”)

- (27) Three aspectual categories of Beja (strong verb, Conj III) with the verb *danbiil* “to collect”

	Perfective	Imperfective	Perfect
Sg. 3	ʔi-dbīl	danbiil	ʔ-ii-dbīl
Sg. 1	ʔa-dbīl	ʔa-danbiil	ʔ-ii-dbīl
Pl. 3	ʔ-i-dbil-‘na	ʔ-ee-dbil-‘na	ʔ-ii-dbil-‘na
Pl. 1	n-i-dbīl	n-ee-dbīl	n-ii-dbīl

As far as the reconstruction of the “marked” category of the perfect is concerned, it is important to realize that at least five different strategies were exploited in its morphological implementation:

- (a) The “be”-perfect in Egyptian is based on the verbal adjective enlarged by the copula *-n* (finitized by oblique forms of pronominal clitics): */sajím-n = Vj/ “I have heard,” */sajím-n = Vk/ “you have heard,” */sajím-n = af/ “he has heard” (see Allen 2014: 245–247 for documented forms).
- (b) The middle *t*-perfect in Akkadian and Eblaite is ambiguous between the active perfect and middle perfective: *i-p-ta-ras* “he has separated” ~ *i-l-ta-baš* “he dressed himself” (see Bubenik 2017: 49–52). The active interpretation was common only with a small group of verbs denoting “holding, grasping, seizing, taking” (Huehnergard 1997: 394). In Classical Arabic the middle semantics is found with some verbs in measure VIII: *samiša* “he (has) heard” > *i-s-ta-maša* “he (has) listened” (called *iptaras*-perfect by Zaborski 2004).
- (c) In Beja strong verbs (Conj III) *danbīl* “he collects” (in 27) the perfect is formed by lengthening the preformative of the perfective: *ʔi-dbīl* “he collected” (perfective) > *ʔii-dbīl* “he has collected” (perfect).
- (d) In Berber the perfect is formed by ablaut: unmarked **ya-knuf*/**ya-knif* “he cooked” > perfect **yu-knaf* “he has cooked” (section “The Role of Ablaut in the Formation of the Three Aspectual Stems”).
- (e) The innovative West Semitic perfect is based on the shift stative > perfect (section “Semitic Parallels in Marking for the Imperfective”). The active “neo”-perfect (*katab-a* “he wrote”) is based on the pattern CaCaC documented with several stative verbs in Old Akkadian and Assyrian. As was shown in section “Stative Conjugation”, the archaic stative conjugation (preserved in Akkadian, Egyptian and Kabyle) was recategorized in Central and South Semitic languages as the perfect conjugated by direct (subject) pronominal suffixes.
- (f) In Classical Arabic and Modern Arabic dialects the innovative perfect can be formed by analytic means of combining the copula with the lexical verb: *kāna (qad) kataba* “he had written” (past perfect), *yakūnu (qad) kataba* “(apparently), he has written” (“evidential” perfect).

Conclusion

The central point in our discussion of the Proto-Berber aspectual system (in 23) was the “trichotomy” of the disyllabic imperfective versus the monosyllabic unmarked and the ablauted monosyllabic perfect. We observed that an earlier state of affairs is seen in the aspectual system reflecting the “dichotomy” of verbs of action versus those of state (in 24). Given the parallel situation in Cushitic and Egyptian we wondered whether it can be projected farther back in time. We observed that the vocalic pattern of the Proto-Berber imperfective corresponds to the same category in Akkadian while the pattern of the unmarked corresponds to the Proto-Semitic perfective/jussive. As far as the reconstruction of the marked category of the perfect is concerned, it is important to realize that at least five different strategies were exploited in its morphological implementation during the history of Afro-Asiatic languages (section “Reconstructing Common Proto-Semito-Berber-Cushitic”). Further progress in this comparative/typological approach to language reconstruction and change requires also revisiting the issue of morphological and semantic markedness of individual aspectual categories

(section “[Berber Aspectual System](#)”). In Afro-Asiatic linguistics there is a theoretical problem of appropriateness of the term “aorist” used for a typologically different category in IE and other languages. In this chapter “unmarked” is used for “aorist” in Berber, corresponding to the marked category of the perfective/jussive in Semitic (called also “preterit” in Akkadian and misleadingly “imperfect tense” in West Semitic). More work should be done on other syntactic parallels of Berber with Semitic where proper identification of aspectual/temporal contrasts plays a fundamental role in any comparative approach to Afro-Asiatic linguistics; namely: analytic verbal formations by means of phrasal verbs and particles (in sections “[Temporal/Aspectual Contrasts](#)” and “[Lexical Aspect Contrasts \(Initial Phase, Repetition of an Event\)](#)”) and conjoining clauses by means of (a) syndetic parataxis (in section “[Conjoining Clauses in Berber and Semitic](#)”).

References

- Allen JP (2014) Middle Egyptian. Cambridge University Press, Cambridge
- Appleyard D (2012) Cushitic. In: Edzard L (ed) Semitic and Afroasiatic: challenges and opportunities. Harrassowitz Verlag, Wiesbaden, pp 131–198
- Binnick R (ed) (2012) The Oxford handbook of tense and aspect. Oxford University Press, Oxford
- Bubenik V (2017) Development of tense/aspect in Semitic in the context of Afro-Asiatic languages. John Benjamins, Amsterdam
- Carstairs-McCarthy A (1992) Current morphology. Routledge, New York/London
- Chaker S (1983) Un parler berbère d’Algérie (Kabylie). Université de Provence, Provence
- Comrie B (1976) Aspect. Cambridge University Press, Cambridge
- Galand L (1987) Les emplois de l’aoriste sans particule en berbère. In: Jungraithmeyr H, Müller W (eds) Proceedings of the fourth international Hamito-Semitic congress. John Benjamins, Amsterdam, pp 361–379
- Hewson J, Bubenik V (1997) Tense and aspect in Indo-European languages. Theory, typology, Diachrony. John Benjamins, Amsterdam
- Huehnergard J (1997) A grammar of Akkadian. Scholars Press, Atlanta
- Kossmann MG (2003) The origin of the Berber ‘participle’. In: Bender ML, Takacs G, Appleyard DL (eds) Selected comparative-historical studies in memoriam of Igor M. Diakonoff. LINCOM EUROPA, München, pp 41–54
- Lipinski E (2001) Semitic languages. Outline of a comparative grammar: Leuven, Peeters
- Mettouchi A (2004) Diathesis, aspect and stativity in Taqbaylit Berber. In: Nait-Zerrad K, Voßen R, Ibrizimow D (eds) Nouvelles études berbères II: Le verbe et autres articles. Rüdiger Köppe Verlag, Köln, pp 95–110
- Penchoen TG (1973) Tamazight of the Ayt Ndhir. Undena Publications, Malibu
- Prasse K-G (1973) Vol.2, Manuel de grammaire touaregue. Akademisk Forlag, Copenhagen
- Prasse K-G (2003) ‘La vocalisation du protoberbere’ in M. Lionel Bender et al. eds. selected comparative-historical afrasian linguistic studies, 41–54. LINCOM EUROPA
- Stroemer H (2002) Tashelhiyt Berber folktales from Tazerwalt (South Morocco). Rüdiger Köppe Verlag, Köln
- Voigt RM (1987) Derivatives und flektives im Semito-hamitischen. In: Jungraithmayr H, Müller W (eds) Proceedings of the 4th international Hamito-Semitic conference. John Benjamins, Amsterdam, pp 85–97
- von Soden W (1952) Grundriss der akkadischen Grammatik. Pontificium Institutum Biblicum, Rome
- Zaborski A (2004) Traces of *iptaras* in Arabic. In: Takács G (ed) Egyptian and Semito-Hamitic (Afro-Asiatic) studies in memoriam W. Wycichl. Brill, Leiden, pp 160–171

Marijn van Putten

Contents

Introduction	307
The Perfective Vocalism	310
Application of the A- Θ -A \rightarrow Vocalism to Quadrisyllabic Stems	311
The Aorist Vocalism	312
The Imperfective Formation	315
The Classification of Middleweight Verbs	316
Conclusion	318
Appendix	319
Class 1 (Aor. Θ -A \rightarrow)	319
Class 2 (Aor. Θ \rightarrow)	326
Bibliography	332

Introduction

The complexities of Berber verbal morphology have long been recognized, and its reconstruction has been undertaken by several authors, most notably by Prasse (1973). In principle we can distinguish between several general verb types, which

In order to not give a false sense of specificity, this chapter uses the regular Berberological transcription in the Proto-Berber reconstructions using z and d for the emphatics, y for the back voiced fricative, ə and ā for the central vowels, and k and g for the palato-velars. Transcriptions of living dialects have adopted the transcription of this volume using IPA. The only deviation I adopt is that I use ā for the low central vowel in dialects that distinguish it from ə rather than the IPA sign $[\text{ɐ}]$, and geminates are indicated with doubled consonants $[\text{dd}]$ rather than $[\text{d:}]$.

M. van Putten (✉)

Leiden University, Leiden, The Netherlands

e-mail: m.van.putten@hum.leidenuniv.nl

Heath (2005) has classified for Tuareg as “light verbs” (verbs with two stem vowels) and “heavy verbs” (verbs with more than two stem vowels).

The most common class of light verbs, which shares many similarities with the Proto-Semitic verbal system, as recently examined by Kossmann and Suchard (2018), was first reconstructed for Proto-Berber by Prasse (1973), and received important updates by Kossmann (2001a), especially integrating the newly discovered Proto-Berber consonant *ʔ, and making the reconstructions somewhat less Tuareg-centric. The heavy verbs, however, have not yet received the same attention. This chapter aims to reconstruct these heavy verbs.

Besides the light and heavy verbs, there is at least one other verb class that seems to be reconstructible for Proto-Berber, namely, the set of (mostly stative) verbs that have an *i- or *u- in the initial syllable of the Aorist and Imperfective, e.g., *ini “to say,” and *uksuḍ “to be afraid.” These verbs will have to be discussed in a later publication.

Unlike other Berber languages, Tuareg, Ghadames, and Zenaga Berber¹ retain a distinction between two central vowels *ǎ and *ə, which have merged in the other Berber languages to a single central vowel ə, of questionable phonemic status. This central vowel distinction plays an important role in the apophony of the different verb forms in the languages that retain the contrast. While there are several in-depth descriptions of the heavy verbs in a variety of different Berber dialects, and cross-dialectally (e.g., Boumalk 2003, for Tashelhiyt; Dallet 1953, for Kabyle; and Basset 1929, for a comparative description), no attempt at the reconstruction of the Proto-Berber heavy verbal system has been undertaken and an integration of especially the Tuareg, Ghadames, and Zenaga Berber data into this comparative perspective has also not been undertaken in great detail.

Tuareg displays an astounding number of “heavy verb types,” which Prasse (1973) classifies into 15 different conjugations, with many different subtypes. However, these conjugations share more similarities with each other than that they do with the basic “light verbs,” and recently several authors have significantly reduced the Tuareg heavy verb categories to a significantly smaller subset. Sudlow (2001) takes the classification of Prasse (1973, which in itself was already a simplification of de Foucauld 1950), and reduces it to five basic heavy stems – a similar reduction created by Heath (2005). Kossmann (2011) reconstructs four apophonic classes, which he further reduces to only two general patterns on a historical level (Table 15.1).

Table 15.1 Tuareg apophony

	Aorist	Perfective
Type 2	Θ→	A-Θ-A→
Type 3	A→	Θ-Θ-A → (< *A-Θ-A→)

¹ In Zenaga Berber the distinction between plain vowels *a, *i, *u, *e, and the central vowels *ǎ and *ə is largely lost, generally merging *a and *ǎ to a single phoneme /a/ and *ə to *i or *u, depending on the phonemic environment.

Kossmann envisions these vowel patterns to be assigned to the vowel slots of a stem from left to right, and every vowel that follows after the defined vowel will copy the preceding vowel. A gives rise to *ǎ* on central vowel slots, and *a* on plain vowel slots; whereas *Θ* gives rise to *ə* in central vowel slots, and *i* or *u* in plain vowel slots. As such, a verb such as *vmvnu²kv²l* “to be a chief” (a type 2 verb) can be assigned an Aorist and Perfective, as follows:

Apophony of *vmvnu²kv²l*

Aorist	Perfective
Θ-Θ-Θ-Θ	A-Θ-A-A
<i>əmənukəl</i>	<i>ǎmnakāl</i> (< * <i>ǎmənakāl</i>)

In Kossmann’s work on Ghadames (Kossmann 2013a: 66–74), a similar apophonic class system is identified (Table 15.2).

The third Berber language that also retains the central vowel contrast between **ǎ* and **ə*, Zenaga, is also identified as having a very similar system by Cohen and Taine-Cheikh (2000: 291ff.). Only one clear type is visible in Zenaga, which corresponds mostly to Tuareg Type 2 and Ghadames Class 2. A smaller subtype has A-type vowels in the middle vowel slot, both in the Aorist, and the Perfective. This subtype has no clear parallels in other Berber varieties, and it is, therefore, difficult to decide where it comes from, and how it relates to the verbal type attested in the other varieties (Table 15.3).

Kossmann (2012: 43f.) already observed several of the similarities in vowel patterns of these heavy verbs across the different Berber languages and this chapter wishes to further expand on their linguistic history. The similarities in these systems, across very widespread Berber varieties, makes it highly likely that this is an old system that may be reconstructed to Proto-Berber. In this chapter, I aim to reconstruct the different apophonic classes present in the Proto-Berber heavy verbs, and will show that the systems reconstructible from these *ə/ǎ* contrast–retaining dialects are corroborated by the dialects that have lost this contrast in the outcome of plain vowels in the stem.

Table 15.2 Ghadames apophony

	Aorist	Future	Perfective
Class 2	Θ→	= Aorist	A-Θ-A→
Class 3	Θ→	Θ-A→	A-Θ-A→
Class 4	Θ-A→	= Aorist	A-Θ-A→

Table 15.3 Zenaga apophony

	Aorist	Perfective
Basic	Θ-Θ-Θ	A-Θ-A
Mid-A:	Θ-A-Θ	A-Θ-A/A-A-A

²Central vowel slots are designated with *v*, whereas plain vowels are generally marked with the plain *a*, *i*, or *u*, as it appears on the Aorist.

Throughout this chapter several verbs will be cited, which return in the appendix. These verbs are marked by {#}, where # is the corresponding number in the appendix. In the appendix, one can also find the abbreviations used for the different Berber varieties, and the sources consulted.

The Perfective Vocalism

The perfective vocalism of the heavy stems is clearly reconstructible from the comparative evidence available to us. Ghadames and Zenaga point to a vocalism A- Θ -A \rightarrow in the perfective, while Tuareg, either has Θ - Θ -A \rightarrow , or A- Θ -A \rightarrow . The former Tuareg pattern can easily be understood as being the result of an active Tuareg-internal vowel harmony that shifts any \check{a} preceding ∂ to ∂ (Kossmann 2011: 57 f.). This is clear from the fact that the third-person masculine plural pronominal object clitic =*tān* undergoes vowel harmony to =*tən* when the following vowel is ∂ , when it is in fronted position – while, when it comes after the verb, it remains =*tān* (Kossmann 2011: 22).

Vowel harmony of \check{a} to ∂ in Tuareg

$\partial kr\partial d = t\check{a}n$	“attach them!”
$a = t\partial n = y\partial kr\partial d$	“he will attach them”

This vowel harmony does not apply when the medial high vowel is *u*: in such cases therefore the initial \check{a} reappears (Kossmann 2011: 57), cf.:

P $\partial qq\partial t\check{a}s$ “be cut” < $*\check{a}qq\partial t\check{a}s$
 P $\check{a}ggol\check{a}z$ “be left over” < $*\check{a}ggul\check{a}z$

As such, the A- Θ -A \rightarrow pattern can be reconstructed as the perfective vocalism of the heavy stems in all three varieties.

Traces of this pattern can also be recovered in Berber varieties that have lost central vowel contrast, e.g., Central Moroccan Berber retains a reflex -*a* of word-final $*-\check{a}?$; whereas $*-\partial?$ has no reflex (Kossmann 2001a). Perfectives of heavy roots with final glottal stop, therefore, consistently have a zero-reflex in the aorist and a final *a* in the perfective.

{4} $*\check{a}qq\partial ym\check{a}?$ “to sit,” CMB *qqima* (A *qqim*); Tu. *āqqim*; Ghd. *āqqim*
 {5} $*\check{a}ff\check{a}st\check{a}?$ “to be silent,” CMB *fəsθa* (A *fəsθ*); Ghd. *āfəss*; Zng. *āffus^sʿa?*

Likewise, nouns that have a plain vowel as the second vowel slot will shift it to the expected high vowel in the stem. The vowel usually becomes *u*, unless a consonant *w* follows it, in which case it regularly becomes *i*.³

³ This same shift is attested in the light verbs, with an initial plain vowel followed by *w*, cf. A $*ag\partial m$ P $*ug\check{a}m$ “to draw water,” but A $*aw\partial d$ P $*iw\check{a}d$ “to arrive.”

- From this evidence, it seems safe to conclude that the perfective vocalism A- \emptyset -A \rightarrow is not only present in Tuareg, Ghadames, and Zenaga, and traces of it can be found in many Berber varieties that have lost the short vowel contrast, as such this vocalism can be reconstructed for the Proto-Berber heavy verb.

{31} Kb. *a friwəs* < **ǣfriwəs* P. *frawəs* < **ǣfərawəs* “to jump up”
 Kb. *friræþ* P. *fraræþ* “to roll; tumble”
 Kb. *flili* P. *flali* “to emerge”
 Kb. *friri* P. *frari* “to glide”
 Kb. *clil* < **ǣklilaþ*? P. *clal* < **ǣkalalā*? “to merit; deserve”

However, also $\bar{c}icc$ stems have this i/a alternation in Kabyle, an environment where Tuareg, or any other Berber language seems to undergo this alternation, e.g., Kb. *zziḩar* P *zzaḩar* “to twist; to make a string”.⁴ This alternation is, presumably, analogically spread from the $ccicc$ verbs discussed above, but it is difficult to be sure.

In most other Berber varieties, this alternation has been leveled out. Tashlhiyt, for example, gives no indication of this alternation:

{30} Tashl. A *griwl* P *griwl* “to turn” not ****grawl**

On the basis of the isolated verbs that have the alternation of Tuareg it seems likely that we can reconstruct this alternation to Proto-Berber for verbs with a stem-internal i after two consonants. However, verbs that have a vowel u in this position hardly ever give evidence for this alternation outside of Tuareg, e.g.,

{22} Kb. *fruri* P *fruri* “to drop fruits or seeds;” Tashl. *fruri* P *fruri* “to drop fruits or seeds;”

CMB *frury* P *frury* “to drop fruits or seeds,” not ****frary**.

{20} Tashl. A *krurri* P *krurri* “to be round,” not ****krarri**

The treatment of verbs with a root-internal u in Ghadames follows the majority of the Berber varieties. While it normally applies the A- Θ -A \rightarrow pattern to heavy verbs in the perfective, for verbs of this type, we find a pattern A- Θ - Θ -A, meaning that the u is not changed to a :

{21} Ghd. *gnunni* < **əkənnunəy* P *āgnunne* < **ākənnunāy* “to roll”

{24} Ghd. *āḡlulu* < **əḡəlulu* P *āḡəlula* “to hang,” Fig. A = P *ylulu* “id.”

Basset (1929: 90, 93) records one Kabyle verb with the alternation as found in Tuareg *grurəz* P *grarəz* “to fall into ruin.” This form is not recorded by Dallet in his dictionary. He also records (optionally) the alternation for several verbs in Ida u Semlal: *mdudu* P 1sg. *mdudaḩ/mdadaḩ* “to have insolation” and *flufu* P *flufa* “to boil,” but plural perfective participle *flafanin*.

While the evidence for u/a alternation in verbs of this type outside of Tuareg is even more sparse than it is for the i/a alternation, the fact that it occasionally shows up in fairly disparate varieties of Berber suggests that Tuareg does reflect the original situation, despite this vocalization being lost in most varieties.

The Aorist Vocalism

The Aorist presents a more complicated situation than the Perfective. It is clear that in Tuareg there are two vocalisms, (1) an A \rightarrow vocalism, and (2) an Θ \rightarrow vocalism (Kossmann 2011: 58).

⁴This verb is clearly deverbal from *izikər* “rope.”

A→: Tu. A *āqqātās* P *āqqātās* “to be cut”

Θ→: Tu. A *āggulāz* P *āggolāz* “to be left over”

Zenaga usually points to an Θ→ vocalism even where most Berber languages point to a vocalism containing A-type vowels. This is likely a secondary development (Cohen and Taine-Cheikh 2000: 292):

{8} Zng. A *y-uruʔri* P *y-āruʔrā* “vomit”; CMB *rar* P *rura* “id.”

{12} Zng. A *y-iyiyi* P *y-āyiyā* “be born”; Tam *lal* P *lula* “id.”

{3} Zng. A *y-āddurkiy*⁵ P *y-āddurkāy* “be blind”; Tu. *āddārkal* P *āddārkal* “id.”

In two cases, an A-A-Θ vocalism surfaces as an alternative vocalization for the final glottal stop verbs, once in the clearly reconstructible verb “to be quiet,” and the Arabic loanword “to pray”:

{5} Zng. A *y-uffūsʹsʹi*, *y-affāsʹsʹi* P *y-āffūsʹsʹa* “be silent”; Ghd. *fəss* A *āfəss*

{15} Zng. A *y-usʹsʹulli*, *y-asʹsʹalli* P *y-asʹsʹəllä* “pray”; CMB *zʹzʹall* P *zʹzʹulla*

It is tempting to see in the A-A-Θ vocalism the A→ vocalism found in Tuareg, but the exact development is not clear.

A few cases exist in Zenaga where the Aorist has a Θ-A-Θ vocalism; in these cases, the perfect has an A-A-A vocalism:

Zng. A *y-inmāræg* P *y-ānmāræg* “to resemble”

Zng. A *y-əfwādi:h* P *y-āfwāda:h* “envoyer”

No traces of this subtype can be found in Tuareg or Ghadames. Both words identified by Cohen and Taine-Cheikh (2000) with this vocalism lack clear Proto-Berber etymologies, it is, therefore, difficult to evaluate whether this lexically determined vocalism is an archaism, or rather, an innovation of Zenaga.

Whatever the exact development of the subtypes, it is clear that the Θ→ vocalism can be readily identified with the identical Θ→ vocalism found in Tuareg (and Ghadames) – and that it has been generalized across many verbs where Tuareg has an A→ vocalism in the Aorist.

The situation is more complex for Ghadames. Unlike Tuareg and Zenaga, Ghadames has three different aspectual forms, which all perform different parts of the functions where Tuareg and Zenaga only use the Aorist, namely, the Imperative, Aorist, and Future.⁶ There are four different combinations. Type 1 has Θ→ for all three aspects. Type 4 has Θ-A(-Θ) for all three aspects. Both, Type 2, and Type 3 have Θ-A→ for the Future, but the Imperative, either agrees with the Aorist Θ→ vocalism (Type 2), or with the Future Θ-A→ vocalism (Type 3).

⁵This must be a loanword from another Berber language, as Proto-Berber **y* yields *ʔ* in Zenaga (Kossmann 2001a: 63–65; Souag 2017).

⁶For an overview of the function of these stems see Kossmann (2013a: 161ff.).

	Type 1	Type 2	Type 3	Type 4
Imperative	(Θ -) Θ - Θ <i>fənzər</i> “have a nosebleed”	Θ - Θ - Θ <i>əmmiɬəf</i> “be entered”	(Θ -)A-A <i>māklaw</i> “take lunch”	(Θ -)A(- Θ) <i>fad</i> “be thirsty”
Aorist	Θ - Θ - Θ <i>əffənzər</i>	Θ - Θ - Θ <i>əmmiɬəf</i>	Θ - Θ - Θ <i>əmməkluw</i>	Θ -A(- Θ) <i>əffad</i>
Future	Θ - Θ - Θ <i>əffənzər</i>	Θ -A-A <i>əmmatāf</i>	Θ -A-A <i>əmmāklaw</i>	Θ -A(- Θ) <i>əffad</i>

Other Berber varieties such as Tuareg have a separate Imperative as well, but they never differ in vocalism, but only clip off the first vowel and shorten the initial geminate.

- Tu. Imp. *ɛātās* A *āqqātās* “to cut”
 {4} Tu. Imp. *ɛāym* A *āqqāym* “to sit”
 Tu. Imp. *dubən* A *əddubən* “to marry”

This same process of clipping and shortening is present in Ghadames as well (Kossmann 2013a: 78):

- {5} Ghd. Imp. *ǰäss* A *yāffäss* “to be quiet”
 {7} Ghd. Imp. *βərgⁱ* A *əββərgⁱ* “to dream”

The Tuareg and Zenaga $\Theta \rightarrow$ vocalism can be readily equated with the Ghadames vocalism $\Theta \rightarrow$ (Type 1). The Tuareg Aorist A \rightarrow should probably be equated with the Ghadames Θ -A \rightarrow vocalism. However, these are not perfectly identical, and the question becomes which of the two vocalisms is original to Proto-Berber.

To my mind, it is more likely that the more complex Θ -A \rightarrow vocalism is original, as it seems difficult to arrive at this vocalism from an A \rightarrow vocalism while the Tuareg A \rightarrow vocalism could be derived analogically from an original Θ -A \rightarrow vocalism through the imperative. As already discussed, the imperative clips off the first vowel of heavy stems, and degeminates the word-initial consonant. This yields an A \rightarrow vocalism for the Imperative in Ghadames as well, by virtue of the initial vowel having been clipped. Tuareg could have introduced an innovative A \rightarrow vocalism to the Aorist from the imperative, which in Tuareg is otherwise identical to the Aorist.

Another argument that speaks in favor of the reconstruction of the Proto-Berber Aorist vocalism as Θ -A \rightarrow is that it is the vocalism used in both Tuareg and Ghadames for the Imperfective. In several other verb types,⁷ the Imperfective takes on the same vocalism as the Aorist, so it seems likely that this would also be the case for the Heavy Verbs.

As such, it seems reasonable to assume that the Aorist vocalism of Tuareg was originally identical to that of the imperfective and, thus, also had an

⁷ Verb types that copy the vocalism of the Aorist for the imperfective are most prominently verbs with initial *i*- or *u*- prefixes, e.g., A *ini* “to say” P *ənna* I *ətini* (Prasse’s I.C verbs); A *irsan* “to be tired” P *ārsin* I *ətirsan* (Prasse’s II verbs), and the heavy verbs with an $\Theta \rightarrow$ vocalism in the Aorist, as we will see in section “[The Imperfective Formation](#)” below.

Θ -A \rightarrow vocalism. Therefore, for Proto-Berber it seems possible to reconstruct two main vocalisms in the Aorist, $\Theta\rightarrow$ and Θ -A \rightarrow , the latter of which usually corresponds to the Ghadames Future stem.

Where Tuareg, and most other Berber varieties have an Aorist, Ghadames has three separate stems: the Imperative, Future, and Aorist. Kossmann (2000) convincingly shows that the Future stem should be reconstructed for Proto-Berber, as in eastern Berber varieties besides Ghadames, namely, Sokni, Foqahi, and Awjili, show traces of an ancient Future stem different from the Aorist in the light verb system. However, due to a lack of data and many innovations, these varieties cannot be used to inform us about the vocalism of Future stem of the heavy verbs. As such, it remains unclear to what extent the system of four types of heavy verbs in Ghadames is ancient, or not. What we can observe is that the two vocalisms reconstructible for the other Berber languages are also present, in some form, in Ghadames. The Type 1 verbs, which have $\Theta\rightarrow$ vocalism for all three stems can be readily equated to the verbs that have an Aorist with $\Theta\rightarrow$ vocalism in other varieties. Types 2, 3, and 4 all have Θ -A \rightarrow vocalism in at least one of the three stems that correspond to the Aorist in other varieties. Therefore, this one likely corresponds to the A \rightarrow class of Tuareg. Nothing more can be said about the reconstructibility of the Future in heavy stems of Ghadames.

To sum up, we find that the Perfective vocalism is always A- Θ -A \rightarrow . But the Aorist vocalism consists of two classes. Class 1 can be reconstructed with an Aorist vocalism Θ -A \rightarrow , and Class 2 with an Aorist vocalism $\Theta\rightarrow$.

The Imperfective Formation

The Imperfective formation in Berber presents us with several problems. Where the Aorist and Perfective are, generally, regularly derivable from the stem, the Imperfective in a variety of different Berber dialects represent significantly more issues. This much is clear, for example, in Kabyle, where a verb very frequently has a variety of different Imperfective formations (see, e.g., Dallet 1953: 375ff.).⁸

Several Berber dialects have more predictable imperfective formations, most notably Tuareg, Tashlhiyt, and Central Moroccan Berber. While it seems likely that there were other Imperfective formations around,⁹ the pattern found in these languages is most readily reconstructible. The process usually employed by Tuareg,

⁸This is also clear in the appendix at the end of this chapter, where Kabyle often lists a variety of different imperfective formations.

⁹These may have even had slightly different aspectual functions. Tarifit, for example, makes use of a variety of different Imperfective stems, which express subtle differences in semantics (Lafkioui 2018: 98). While Lafkioui, convincingly, makes the case that in its formal application, this is an innovation of Tarifit, the amount of different stems different verbs can have in Kabyle (e.g., *ɛwəl* ‘to falsify,’ Imperfective *idəwəl*, *yətsiṣədwal*, *yəddəwəl*) suggests that there may have already been some morphological variation which, perhaps, had some semantic specification at an earlier stage of the language already.

However, for quadriconsonantal verbs, the evidence is conflicted. Tuareg, Tashelhiyt, and Central Moroccan Berber generally treat them as a class 1 verbs, whereas Kabyle and Zenatic varieties¹⁰ treat them as class 2 verbs, for example,

- {3} Tuareg *dārɣāl* I *əttadārɣal* “to be blind”; Tashl. *ddrɣl* I *ttɣrɣal* “id.”; CMB *ddərɣəl* I *ttōərɣəl* “id.”;
Kb. *ddərɣəl* I *ttōərɣil* “id.”; Fig. *tʔərɣəl* I *ttʔərɣil* “id.”

However, there is a fairly large group of common verbs for which the majority of the dialects treat them as Class 1, for example,

- {4} **əqqāymā?* I **əttāyāyma?* “to sit; stand still”
{5} **əffāstā?* I **əttāfāsta?* “to be silent”
{8} **əṛā?ṛā?* I **əttāṛā?ṛa?* “to return; vomit”
{10} **əḡḡawān* I **əttāyawan* “to be full; satiated”
{12} **əlālā?* I **əttālala?* “to give birth”

Therefore, it is tempting to see the treatment of Kabyle and Zenatic of verbs like {3} **əddāryāl*, as Class 2 verbs as innovations, rather than retentions. With the loss of central vowel contrast, which took place in both varieties, the distinction between Class 1 and Class 2 verbs would have gotten blurred, and therefore, it stands to reason that some varieties would have analogically spread the Class 2 imperfective formation to verbs that were originally part of Class 1.

There is one group of quadriconsonantal verbs, however, which have Class 2 vocalism in the imperfective, even in Tashelhiyt and Central Moroccan Berber (but not in Tuareg, where they are Class 1). These are the verbs that have complete reduplication of two root consonants.

- {25} Kb. *ffərɣər* I *tɣərɣir* “to flap the wings”; CMB *fərɣər* I *tɣərɣir* “id.”; Tashl. *ffərɣər* I *tɣərɣir* “id.”
{26} Tu *bārbār* I *əttabārbar* “to go out”; Ouargla *bbərɣər* I *əttbərɣir* “to go up; to travel”
{28} Tu *dāḡdāḡ* I *əttadāḡdag* “to crush”; Kb. *ddəjɔj* I *ddəjɔj* “id.”; Ayt Atta CMB *dəḡdəḡ* I *ddəḡdḡ* “id.”

The fact that, otherwise rather diverse Berber varieties are in agreement that such verbs are Class 2 verbs – except for Tuareg – makes it likely that Tuareg is innovative in this regard, and that reduplicated verbs belong to the Class 2 verbs, despite their only having three central vowels in the stem, which is a feature that would normally qualify them to belong to the middleweight Class 1 verbs.

A confounding element of this reconstruction, however, is that Figuig shows special treatment of these reduplicating verbs that is different from both Class 1 and Class 2 verbs. Unreduplicated verbs in Figuig have the imperfective *tcəccic*, consistent with Class 2 (e.g., *ddərɣəl* I *dʔərɣil* “to be blind”); But the reduplicated verbs

¹⁰Zenatic languages form a fairly homogeneous subgroup that share several morphological innovations with each other. For a discussion of these features see Kossmann (1999: 31f.; 2017).

are regularly *tcəccəc* (e.g., *ffərɣər* I *tfərɣər* “to flap the wings”) – with no long vowel infixation, whatsoever (Kossmann 1997: 139). This appears to be a behavior unique to Fighu as other Zenatic varieties, such as Ouargli, give such verbs *i*-infixes. At the moment it is unclear how the Fighu data should be integrated into our understanding of the reconstruction of the heavy verb.

Conclusion

In this article we have shown that the heavy verb classes can be reconstructed for Proto-Berber in the form of two distinct apophonic classes. The apophonic class to which any one heavy verb belongs is largely predicted by its stem shape. Class 1 is the class of verbs that contain three central vowels in the stem, and a few lexically determined stems with a plain vowel (see {9–13}).¹¹ Class 2 contains the majority of the verbs with plain vowels, verbs with four central vowels, and a small group of reduplicated verbs with three central vowels. Schematically the vocalizations can be represented as per Table 15.4.

There still remain several verb classes that are clearly reconstructible for Proto-Berber. First, there is a group of verbs with a vocalic prefix in the Aorist, such as Tashl. A *ini* P *nn^a/i* I *ttini* “to say”; Kb. A *izmir* P *uzmar* I *tst̪s̪izmir* “to be able”; Tu. A *uksad^ε* P *āksud^ε* I *ttuksad^ε* “to be afraid,” and second there are the causative verbs derived from, either light or heavy verbs, which form a subtype of the heavy verb class. Besides this, there are the reciprocal derivations with an *m*-prefix, and passive derivations with a *t*-prefix, which seem to have a vocalic pattern that agrees with Class 1 verbs. In Tuareg and Ghadamsi, they simply behave as Class 1 verbs, but in other Berber languages they often receive an *a*-infix in their stem. A full treatment of these forms will have to be discussed another time.

Lexical Data

Throughout this chapter, lexical data from a variety of different Berber dialects have been used. The following table gives the abbreviation used, name of dialect, and reference to the source.

Table 15.4 Proto-Berber apophony

	Aorist = Imperfective	Perfective
Class 1	Θ-A→	A-Θ-A→
Class 2	Θ→	A-Θ-A→

¹¹ While the subset of non-derived verbs with Class 1 vocalism is smaller than those with Class 2, verbs that are in this class, they are generally easier to reconstruct for Proto-Berber than those in Class 2. The Class 1 vocalism is also used for Medio-passives (*m*-derivations) and passives (*t*-derivations) in Tuareg, and traces of this can be found in other Berber languages as well. However, several other formations seem to be common in these derivational patterns that do not easily map onto Class 1 verbs in other Berber languages, such as the Tashl. A *myissin* P *myassan* I *tmyissin* “to know one another” (Boumalk 2003: 128).

Abbreviation	Dialect name	Source
Kb.	Kabyle	Dallet (1982)
CMB	Central Moroccan Berber	Taifi (1992); Oussikoum (2013)
Tashl.	Tashelhiyt	Stroomer (forthcoming)
Ghd.	Ghadames	Lanfry (1973)
Tu. M	Mali Tuareg	Heath (2006)
Tu. U	Tudalt Tuareg (Burkina Faso)	Sudlow (2009)
TU. BF-M	Tamaghit Tuareg (Burkina Faso)	Sudlow (2009)
Tu. W	Iwellemeden Tuareg (Niger)	Prasse et al. (1998)
Tu. Y	Ayer Tuareg (Niger)	Prasse et al. (1998)
Fig.	Figuig	Benamara (2013), and Kossmann (1997)
Zng.	Zenaga	Taine-Cheikh (2008, 2010)
Awj.	Awjila	Van Putten (2014)
Snh.	Senhaja	Ibañez (1959)
Rif.	Tarifit	Serhoual (2002)

Appendix

This appendix gives an overview of reconstructible heavy verbs of the different types. It does not intend to be exhaustive, but it does attempt to give several clear and illustrative examples of the different formations that exist. These examples are numbered, and whenever verb forms are referred to in the article, the corresponding number is given. In several cases, previous works have commented on the reconstruction, or cognates present, of some of these reconstructed words. References to these are given in square brackets after the reconstructed term. The etymologies discussed in Kossmann (1999) are marked by K, followed by the example number from the book. Cognates discussed in Naït-Zerrad (1998–2002) are marked by NZ, followed by the abstracted root as found in that work. The etymological notes of Taine-Cheikh (2008) are marked by TC, followed by the footnote number that discusses the etymology. Finally, if the verb is discussed by Basset (1929) it is marked by B, followed by the conjugation number where it is discussed.

Class 1 (Aor. Θ -A→)

* əcācāc

This group is rather common in Tuareg but is significantly more difficult to detect in other Berber languages. When we do find these verbs, we usually find them as transitive āccāc /intransitive əcācāc pairs, much like Tuareg ābtas “to cut,” bātās “to be cut.” None of the verbs of this type that I have detected in northern Berber have good cognates in Tuareg. While this pattern is not productive in any modern Berber variety, it likely points to an older situation where this derivational pattern was more productive.

{1} *əqqārās “to be torn”

This verb corresponds to the transitive light verb *āyrās “to slaughter; to tear.”

Proto-Berber	Aorist *əqqārās	Perfective *əqqārās	Imperfective *əttāyāras	
Kabyle	qqər ^ʰ s ^ʰ	qqər ^ʰ s ^ʰ	ṭṣṭəkras	“to be torn”
Figuig	qr ^ʰ əs ^ʰ	qr ^ʰ əs ^ʰ	[qər ^ʰ r ^ʰ əs ^ʰ]	“to be torn; tear”

The Figuig form has been reanalyzed as a regular light verb, but the retention of the *q*, rather than *ʰ* in initial position shows that it was originally derived from a verb with an initial *qq*.

In Kabyle we still find quite a lot of verb pairs of this type, but most of these lack well-attested cognates in other Berber varieties, or, when there are cognates, only show cognates for one of the two forms of the pair. Not all of them have as clear a distinction between transitive and intransitive. Some examples are: *llfəʰ* ~ *lfəʰ* “to be destroyed,” *lləʰs* “to be very wet” ~ *lʰəs* “to be wet.” One should perhaps also compare Kabyle *nnkəl* “to spill; to be spilled” and Tuareg (M) *əḡkəl* “to be poured out.” This is likely an ancient retention of Kabyle, but without more widespread comparative evidence it is difficult to be sure.

{2} *ərrāzā? “to be broken”

This verb corresponds to *ārə? “to break.”

Proto-Berber	Aorist *ərrāzā?	Perfective *ārrəzā?	Imperfective *əttārāza?	
Kabyle	r ^ʰ r ^ʰ əz ^ʰ	r ^ʰ r ^ʰ əz ^ʰ	ṭṣṭər ^ʰ z ^ʰ a, [ṭsr ^ʰ uz ^ʰ (u(y))]	“to be broken”
Figuig	r ^ʰ r ^ʰ əz	r ^ʰ r ^ʰ əz	[ttr ^ʰ əddza]	“id.”
Ouargla	r ^ʰ r ^ʰ əz ^ʰ	r ^ʰ r ^ʰ əz ^ʰ	yəttər ^ʰ z ^ʰ a, [yət ^ʰ r ^ʰ ər ^ʰ əz ^ʰ , r ^ʰ r ^ʰ əz ^ʰ]	“id.”

***əċċāċċ**

This stem type has several roots that are readily reconstructible for Proto-Berber with widespread attestations, but the most readily reconstructible forms all end in a final glottal stop.

{3} *əddāryāl “to be blind” [B: 123; NZ: DRYL; TC: 209]

Proto-Berber	Aorist *əddāryāl?	Perfective *əddāryāl?	Imperfective *əttāddāryal?	
Kabyle	ddərəkəl	ddərəkəl	[ddərəkəl, ṭṭərəkəl]	“to be blind”
Tuareg (M)	d ^ʰ ārəkāl	[d ^ʰ ārəkāl-]	ətad ^ʰ ārəkāl	“id.”
CMB	ðər ^ʰ kəl	ðər ^ʰ kəl	[ttðər ^ʰ kəl]	“id.”
Figuig	t ^ʰ t ^ʰ ər ^ʰ kəl	t ^ʰ t ^ʰ ər ^ʰ kəl	[td ^ʰ ər ^ʰ kəl]	“id.”
Ouargla	ddərəkəl	ddərəkəl	[ttddərəkəl]	“id.”
Tashlhiyt	drəkəl	drəkəl	ttddərəkəl	“id.”
Zenaga	əddurkiy	əddurkiy	[əddurkiy]	“to be one-eyed”

The Zenaga form is not cognate, but rather a loan from another Berber language, this is clear because Proto-Berber *ʏ regularly shifts to ʔ in Zenaga, and not to ʁ that we see here (Kossmann 2001a: 63–5; Souag 2017).

Most Berber languages have a $\Theta \rightarrow$ vocalism in the Aorist/Imperfect of this verb. As discussed in the section called “The Classification of Middleweight Verbs”, it seems likely that Tuareg and Central Moroccan Berber retain the original vocalism here, although this could certainly be debated.

One is tempted to see this root as some form of compound of an unidentified element *dr*¹², followed by a shortened form of the verb **āqqəl* “to watch” (cf. Fig. *qqəl* “to watch,” CMB *qqəl* “id.,” and probably Tu. *ixal* “to believe”).¹³

{4} **āqqāymā?* “to sit; remain” [B: 198; TC: 52]

Proto-Berber	Aorist * <i>āqqāymā?</i>	Perfective * <i>āqqəymā?</i>	Imperfective * <i>əttāyāyma?</i>	
Kabyle	<i>qqim</i>	<i>qqim</i>	[<i>ṭṣimi</i> , <i>ṭṣim</i> , <i>ṭṣama</i>]	“to stay; remain”
Tuareg (M)	<i>āqqam</i> , <i>āqqaym</i>	<i>āqqima</i>	<i>ətakāyma</i>	“to sit; be seated”
CMB	<i>qqim</i>	<i>qqima</i>	<i>ttṣima</i>	“to sit; remain”
Figuig	<i>qqim</i> , <i>qqyəm</i>	<i>qqim</i> , <i>qqyəm</i>	<i>ttṣima</i>	“id.”
Ouargla	<i>əqqim</i>	<i>əqqim</i>	<i>əttqima</i>	“id.”
Ghadames (irregular)	Imp. <i>qem</i> < * <i>ṭāymā?</i> [Aor. <i>əqqim</i> < * <i>āqqəymā?</i>]	<i>āqqim</i>	[<i>əttṣim</i>]	“to remain”
Tashlhiyt	[<i>qqama</i> , <i>ṣama</i> , <i>ṣumu</i> , <i>qumu</i>]	[<i>yama</i> , <i>qama</i>]	[<i>ṭṣumu</i> , <i>ttqumu</i> , <i>ttqama</i> , <i>ttṣama</i>]	“to sit; remain”
Zenaga	[<i>iʔmi</i>]	<i>aʔma</i>	[<i>ittiʔmi</i>]	“to sit”

The Tashlhiyt forms are certainly cognate to other varieties, but it is difficult to reconcile its form with the one attested in the other varieties examined here, even when compared to Central Moroccan Berber, which in many other ways is linguistically very close to Tashlhiyt. I do not have a clear historical explanation for the presumably innovative development of the Tashlhiyt form.

{5} **əffāstā?* “be silent” [NZ: FST1; TC: 314]

Proto-Berber	Aorist * <i>əffāstā?</i>	Perfective * <i>əffəstā?</i>	Imperfective * <i>əttāfāsta?</i>	
CMB	<i>fəsθ</i>	<i>fəsθa</i>	<i>təfsθa</i>	“to be silent”
Tashlhiyt	<i>fəss</i> , [fiss]	<i>fəss</i> , [fiss]	<i>tfəssa</i> , [tfssas]	“id.”

(continued)

¹² I have found some marginal evidence that *dr* may have originally been a kind of pejorative prefix, cf. Kb. *ddərçəl* “to misstep,” Kb. *tiçli* “walk,” Tu. M *akəl* “to step.” Perhaps we can add here, also, Mzab *ddərnəs* “to languish,” Mzab *əns* “to spend the night,” Kb. *ddərβəz* “to fall heavily,” Tu. WY *bāzbāz* “to fall (of fruit or grain)”.

¹³ I thank Maarten Kossmann for suggesting this connection to me.

Proto-Berber	Aorist <i>*əffästā?</i>	Perfective <i>*äffästā?</i>	Imperfective <i>*əttäfästa?</i>	
Ghadames	Imp. <i>fäss</i>	<i>yäffäss</i>	<i>əttäfäss</i>	“id.”
Zenaga	<i>[uffus^ʕs^ʕi, äffäs^ʕs^ʕi]</i>	<i>äffus^ʕs^ʕa</i>	<i>[yətʃ^ʕus^ʕs^ʕi]</i>	“id.”

Most Berber varieties have assimilated the final **st* cluster to *ss*, but it seems reasonable to take the Central Moroccan Berber form as original here. The emphatic *s^ʕs^ʕ* in the Zenaga form is not easily explained.

{6} **əggärβä?* “to be last” [K: 140; NZ: GRB1, GR45; TC: 357]

Proto-Berber	Aorist <i>*əggärβä?</i>	Perfective <i>*äggärβä?</i>	Imperfective <i>*əttägärβa?</i>	
CMB	<i>ggru</i> <i>ggir, g^wər</i>	<i>ggra</i> <i>ggir, g^wər</i>	<i>[ttəjru]</i> <i>tgira, tggir, tgg^wər</i>	“to be last”
Tashlhiyt	<i>gg^wru</i>	<i>ggra</i>	<i>[ttg^wru, ttggru, ttgru]</i>	“to be last”
Kabyle	<i>j^wri</i>	<i>j^wra</i>	<i>[tstsj^wray]</i>	
Ghadames	<i>g'ərβ</i>	<i>äg'g'ərβ(o)</i>	<i>ittäg'ərβ</i>	“to be behind”
Zenaga	–	<i>yäggärä</i>	–	“to end up . . .”
Tuareg (M)	<i>[əg'g'urh]</i>	<i>[əg'g'urha]</i>	<i>[tig'ərhu]</i>	“to end up”

The final *u* of Tashlhiyt and some forms of Central Moroccan Berber are perhaps the regular outcome of the sequence **βv?* in word-final position, but could also be an analogical spreading of the final *u* that we find in the nominal *m*-derivation *anggaru* “last” < **a-näggärβu?*

Tuareg has a different derivation, with *u* after the first root consonant, and thus belongs to Class 2 verbs.

{7} **əβa/ärgä?* “to dream” [K: 320; B: 94]

Proto-Berber	Aorist <i>*əβa/ärgä?</i>	Perfective <i>*äβu/ärgä?</i>	Imperfective <i>*əttäβa/ärga?</i>	
Tuareg (M)	<i>[əhhurg'ə-t]</i>	<i>əhhurg'ä-t</i>	<i>[tiħərg'ü-t]</i>	“to dream”
Tuareg (BF-U)	<i>[əhhärg'ä-t]</i>	<i>əhhərg'ä-t</i>	<i>[ti-härg'ä-t]</i>	“id.”
CMB	<i>warja</i>	<i>warja</i>	<i>ttwarja</i>	“id.”
Tashlhiyt	<i>wwarg</i>	<i>wwarg</i>	<i>ttwarga</i>	“id.”
Kabyle	<i>arju</i>	<i>urja</i>	<i>tstsarju</i>	“id.”
Ghadames	<i>əββärgⁱ</i>	<i>yäββärgⁱ</i>	<i>əttäβärgⁱ</i>	“id.”
Awjila	<i>[uf tavərgát]</i>	?	<i>vərga</i>	“id.”
Tarifit	<i>arza</i>	<i>urza</i>	<i>tarzi, tarza</i>	“id.”
Figuig	<i>rzit</i>	<i>rzit</i>	<i>ttiḡrit</i>	“id.”

As with the previous verb, Mali Tuareg has shifted this verb to another class, with *u* after the first root consonant, thus becoming a class 2 verb. The exact origin of the -*t* suffix of Tuareg is not yet fully understood, but often occurs on vowel-final/glottal stop-final verbs.

Central Moroccan Berber and Tashlhiyt – and perhaps Kabyle – point to a stem with **a* in the root **əβārgǎʔ*. The Mali Tuareg form could originally be from the same form. As Class 1 and Class 2 merge in the perfect, it is easy to imagine that a verb like this could shift class by analogy from the perfect. The Tudalt Tuareg form, however points to **əβārgǎʔ*, which likewise seems to be the basis for the Figuig form, which, moreover, has a final **t* that seems to correspond to the Tuareg forms with final *-t*, rather than a final **ʔ*. Awjili and Ghadames seem to point to a short vowel in the stem as well. These disparate forms are not easily reconciled.

{8} **ərǎʔrǎʔ* “to return; vomit” [TC: 787]

Proto-Berber	Aorist <i>*ərǎʔrǎʔ</i>	Perfective <i>*ǎrǎʔrǎʔ</i>	Imperfective <i>*əttǎrǎʔraʔ</i>	
CMB	<i>rar</i>	<i>rura, rur</i>	<i>ttrara</i>	“to return; vomit”
Tashlhiyt	<i>rar</i>	<i>rur</i>	<i>ttrara</i>	“id.”
Kabyle	<i>ərr</i>	<i>ərɾa</i>	<i>[tst̪sarra]</i>	“id.”
Ghadames	<i>ǎrr</i>	<i>ǎrro</i>	<i>ttǎrr</i>	“id.”
Figuig	<i>rr</i>	<i>ərru</i>	<i>tərɾa</i>	“id.”
Ouargla	<i>ərr</i>	<i>ərru</i>	<i>əttərɾa</i>	“id.”
Zenaga	<i>[uruʔri]</i>	<i>ǎruʔrǎ</i>	<i>[ətruʔri]</i>	“id.”

This reduplicated verb does not seem to have the initial gemination found in other verbs of this type. This is perhaps to be related to the general avoidance of gemination of certain “weak,” and highly sonorous, consonants such as **w*, **y*, **β* as described by Kossmann (2008: 283ff.), and **r*, **l*, **m*, **n* as seen in imperfective formations of light verbs in Tashlhiyt and Ghomara, which have *ɛrs* “to slaughter” Impf. *qqrs*, rather than the more generally attested *ɛərrəs* (Dell and Elmedlaoui 1991: 85f.; Mourigh 2016: 143f.). Whatever the explanation, the seemingly likewise (partially) reduplicated verb {11} *əlalǎʔ* “to give birth” also lacks the gemination. Ghadames, Zenatic, and Kabyle lack the full vowel *a/u* in the root that we see in Central Moroccan Berber and Tashlhiyt. This is presumably because of the sequence **əʔ* lost completely in these varieties (and **ǎʔ* when unaccented) (Kossmann 2001a: 91f.). The regular outcome of the verb, e.g., in Kabyle, would therefore have presumably been A **rar* P **ərɾa* I **tst̪srara*. The predicted apophony has been analogically leveled out in all of these varieties.

{9} **əGGǎʔfǎy* “to go up” [NZ: GFY; TC: 331]

Proto-Berber	Aorist <i>*əGGǎʔfǎy</i>	Perfective <i>*ǎGGǎʔfǎy</i>	Imperfective <i>*əttǎGGǎʔfay</i>	
CMB	<i>ggafy</i>	<i>ggufy</i>	<i>tjafay</i>	“to go up”
Zenaga	–	<i>äggufǎh (-a:n)</i>	–	“to head north”

This etymology was suggested already by Taine-Cheikh (2008: 174, note 331) and Kossmann (2018: 181). The last of these seems quite uncertain about the etymology, but the formal match seems fine. Considering the context of where

Zenaga is situated, and was situated, in past centuries, south of the Atlas mountains, it strikes me as quite probable that, indeed, these two words are cognates.

*əcacăc

{10} *əyyawăn (or *əyyăḏwăn) “to be full, satiated” [K707, 711; B: 193; NZ: GWN1]

Proto-Berber	Aorist *əyyawăn	Perfective *əyyiwăn	Imperfective *əttăyawan	
Figuig	dʒawən	dʒiwən	tyawan, dʒawan	“to be satiated”
Ouargla	ʒʒawən	əʒʒiwən	əttʒawan	“id.”
Tarifit	ʒʒawən	ʒʒiwən	tyawan	“id.”
Tuareg (M)	əyyăwăn	əyyəwăn	ətayăwan	“id.”
Tashlhiyt	ʒʒawn, ʒʒiwn	[ʒʒawn]	ttʒʒawan, ttʒiwin	“id.”
CMB	ḏʒḏʒawn	ḏʒḏʒiwn	tyawan	“id.”
Ghadames	(əg'g'iwăn) Fut. əg'g'ewăn ¹⁴	(əg'g'ewăn)	əttăg'ewăn	“id.”
Awjili	iwən		[ttiwăn]	“id.”

Tuareg has a central vowel as its second vowel, and, therefore, behaves just like verbs of the type {1} *əqqrăś “to be ripped.” This same reflex is found in the next verb of the same structure {10} *əGGawăr “to sit.” There is no clear explanation for this distribution, and other verbs in similar environments retain the *a*, e.g., *hawăl* “to speak; rant,” while others seem to vascillate between dialects in terms of their treatment of this sequence, e.g., BF-M *kawăn* “to answer,” but BF-U *kăwăn* “id.” In Ayer Tuareg and the closely related Tudalt dialect of Burkina Faso the *cacăc* class has disappeared completely and thus their treatment of such verbs in this way is to be expected (Sudlow 2001: 206; Prasse et al. 1998: 452).

Kossmann (1999: 232f.) argues convincingly that the alternation between initial *ḏʒḏʒ/ʒʒ* in the Aorist and Perfective and *tt-y* in the imperfective in several of the varieties discussed here is best explained as *ḏʒḏʒ* being the outcome of geminate *yy, as verbs of this type would regularly geminate the initial in the Aorist and Perfective stem. It appears that Awjili and Tuareg have a different outcome of *yy than the other Berber varieties.

{11} *əGGawăr (or *əGGăḏwăr) “to sit” [NZ: GWR1]

Like the previous word, this word has a central vowel in Tuareg. As there are no dialects that distinguish *g from *ġ that retain this root, it is reconstructed here with an undefined voiced velar *G.

¹⁴ One wonders if this form the result of mid-vowel harmony, as discussed in the noun by Van Putten (2018), reflecting *əġġawăn. There is, however, no clear explanation as to why this development would apply here and not to other verbs.

Proto-Berber	Aorist *əGGawār	Perfective *əGGiwār	Imperfective *əttāGawar	
Tuareg (WY)	ǎggǎwār	ǎggəwār	əttǎǎgāwar	“to sit down”
Tashlhiyt	ggawr, [ggiwr]	[ggawr], ggiwr	ttgawar	“id.”
CMB	[jiwr]	jiwr	[tjiwir]	“id.”

{12} *əlalā? “to give birth” [TC: 1074]

Proto-Berber	Aorist *əlalā?	Perfective *ālulā?	Imperfective *əttālala?	
CMB	lal	lula	ttlala	“to give birth”
Tashlhiyt	lal	lul, lula	ttlala	“to be born”
Kabyle	lal	lul	ṭslala, [ṭslal]	“id.”
Ouargla	llal	llul	ttlala	“id.”
Ghadames		yālul “he is born”		“id.”
Zenaga	[iyiyi]	äyivä	əttiyyi	“id.”

In most varieties this verb lacks the initial gemination, see {8} *ərā?rā? “to give back; vomit,” for a discussion.

*əcācāc

{13} *əǧǧallā? “to swear” [K: 440; B: 165; NZ: GL23]

Proto-Berber	Aorist *əǧǧallā?	Perfective *əǧǧullā?	Imperfective *əttǧǧalla?	
CMB	ggall	ggula	ttjalla	“to swear”
Tashlhiyt	ggall	ggull, ggulla	ttgalla	“id.”
Kabyle	ggall	ggull	ṭsjalla	“id.”
Figuig	dʒall	dʒull	dʒalla < *tʒalla	“id.”
Ouargla	ʒʒall	ʒʒull	ttʒalla	“id.”
Tarifit	ʒʒadʒ	ʒʒudʒ	tʒadʒa	“id.”

{14} *əǧǧaggā? “to break up camp” [K: 446; B: 166; NZ: G26]

Proto-Berber	Aorist *əǧǧaggā?	Perfective *əǧǧuggā?	Imperfective *əttǧǧagga?	
CMB	ggadʒdʒ	[ggadʒdʒ]	tjadʒdʒa	“to decamp”
Kabyle	ggadʒ	ggudʒ	ṭsjadʒ	“id.”
Tuareg (M)	ǎǧǧagǧǧ	ǎǧǧugǧa	tagǧagǧa	“to load”

This root underwent a typical dissimilation of two subsequent velars where the latter turned into a $\bar{d}\bar{z}$, or \bar{z} , in Central Moroccan Berber and Kabyle. This is the outcome of the sound law *g/y. . . *g > *g/y. . . *z as described by Kossmann (1999: 228), cf. also {16} *əǧǧuǧal “to be an orphan”.

{15} *əzzallā?* “to pray” [B: 165]**

Proto-Berber	Aorist <i>*əzzallā?</i>	Perfective <i>*əzzullā?</i>	Imperfective <i>*əttāzalla?</i>	
CMB	<i>zʷalʷ</i>	<i>zʷulʷa</i>	<i>dzʷalʷa</i>	“to pray”
Tashlhiyt	<i>zʷall</i>	<i>zʷull, zʷulla</i>	<i>tzʷalla</i>	“id.”
Kabyle	<i>zʷall</i>	<i>zʷull</i>	<i>ṭszʷalla, dʷzʷalla</i>	“id.”
Figuig	<i>zʷall</i>	<i>zʷull</i>	<i>tzʷalla</i>	“id.”
Ouargla	<i>zʷall</i>	<i>zʷull</i>	<i>ttzʷalla</i>	“id.”
Zenaga	<i>yusʷsʷulli, yasʷsʷalli</i>	<i>yasʷsʷəllā</i>	<i>yətsʷallā</i>	“id.”

This verb certainly is a very early loanword from Arabic *sʷallā* “to pray” (Kossmann 2013b: 82) and so cannot be reconstructed for Proto-Berber. It is, however, the only verb of this type besides {13} **əggallā?* “to swear,” well-attested across all Berber languages. Van den Boogert and Kossmann (1997: 318) suggest that this loanword was probably adapted into the class by analogy with {13} **əggallā?* “to swear.” It is given here because it is a good illustration of this verb type in Zenaga, which lacks a cognate of {13} **əggallā?* “to swear.” However, the Zenaga verb has *sʷsʷ* in place of *zʷzʷ*. This is not the regular outcome of Proto-Berber **zz*, which normally yields *zʷzʷ* while simplex **z* yields *θʷ*, e.g., P *yarʷθʷa* I *yirʷazʷzʷa* “to break” (Taine-Cheikh 1999: 317; 2008: 453 f.). The reflex found here is presumably due to the influence of the Arabic verb *sʷallā*, from which it ultimately derives.

Class 2 (Aor. $\Theta \rightarrow$)

Verbs of this type generally have a lexical vowel **u* as their second stem vowel. Those that do not usually have four central vowels.

****əcucəc***

While this type is fairly well-attested, there are rather few reconstructible verbs in this type. Nevertheless, the actual formation is uncontroversial, forming A **əcucəc* P **ăcucăc* I **əttacucuc*.

{16} **əggugəl* “be an orphan” [K: 694; B: 95; NZ: GL22]

Proto-Berber	Aorist <i>*əggugəl</i>	Perfective <i>*əggugəl</i>	Imperfective <i>*əttəgugul</i>	
Tuareg (WY)	<i>əggugəl</i>	<i>əggugəl</i>	<i>əttəgugul</i>	“to be an orphan”
Kabyle	<i>gguzəl</i>	<i>gguzəl</i>	<i>ṭsjuzəl</i>	“id.”
Ouargla	<i>gguzəl</i>	<i>gguzəl</i>	<i>ttguzul</i>	“id.”

A dissimilatory sound law **g/γ...g > *g/γ...ž* (Kossmann 1999: 228) has affected a variety of different dialects, and it cuts right across the Tuareg dialect continuum, where some dialects have not undergone this dissimilation, while others have.

{17} **əffudəy* “be injured” [B: 93; NZ: FDY]

Proto-Berber	Aorist * <i>əffudəy</i>	Perfective * <i>əffudāy</i>	Imperfective * <i>əttəfuduy</i>	
Tuareg (M)	[<i>əffadāy</i>]	<i>əffidāy</i>	[<i>tafaday</i>]	“to have cuts on skin”
Tuareg (WY)	<i>əffudi</i>	<i>əffodāy</i>	<i>ətəfuduy</i>	“to have a sore on the back”
Kabyle	<i>ffuði</i>	<i>ffuði</i>	[<i>tfuði</i>]	“to be injured”
Tashlhiyt	<i>fudi</i>	<i>fudi</i>		“id.”
CMB	<i>ffiðy</i>	<i>ffiðy</i>	<i>ttfiðiy</i>	“id.”

In Mali Tuareg, this verb is a Class 1 verbs *fadāy*, perhaps the result of a reclassification due to Class 1 and Class 2 verbs of this type looking identical in the perfective. Central Moroccan Berber has a *i* vowel in the root, which does not have a clear explanation.

**əcuccəc*{18} **əBBunzər* “have a nosebleed” [B: 168; NZ: BNZR; TC: 559]

Proto-Berber	Aorist * <i>əBBunzər</i>	Perfective * <i>əBBunzār</i>	Imperfective * <i>əttəBunzur</i>	
Tuareg (M)	<i>əhhunfər</i>	<i>əhhunfār</i>	<i>ətihənfur</i>	“have a nose bleed”
CMB	<i>junzər</i>	<i>junzər</i>	<i>tjunzur</i>	“id.”
Tashlhiyt	<i>ggunzr</i> (b-, m-, w-)	<i>ggunzr</i>	<i>ttggunzur</i>	“id.”
Figuig	<i>ppunzər</i>	<i>ppunzər</i>	<i>tbunzur</i>	“id.”
Ouargla	<i>mmunzər</i>	<i>mmunzər</i>	<i>ttmunzur</i>	“id.”
Kabyle	<i>ffunzər</i>	<i>ffunzər</i>	<i>tsfunzur</i>	“id.”
Zenaga	<i>uːnzər</i>	<i>āwnzər</i>	<i>əttuːnzər</i>	“Id.”
Ghadames	<i>əffənzər</i>	<i>əffənzār</i>	<i>əttəfənzər</i>	“id.”

The most well-attested verb in this class is “to have a nosebleed,” which has a rather astounding amount of variation of its initial consonant, which appears to reflect any possible initial labial consonant across the different Berber languages (denoted as a **B* here), assuming here that the *g/gg* reflexes of CMB and Tashlhiyt are the result of geminated **ww* > **gg*^w. Despite this difficulty, the verb is one of the best attested forms of **əcuccəc* verbs and is, therefore, included here.

**əcūcəc*{19} **əddukkəl* “to walk together, accompany” [NZ: DKL5]

Proto-Berber	Aorist * <i>əddukkəl</i>	Perfective * <i>əddukkāl</i>	Imperfective * <i>əttədukkul</i>	
Tuareg (B-M)	<i>əddukkəl</i>	<i>ədokkāl</i>	<i>ətidukkul</i>	“gather together”
Tashlhiyt	<i>ddukkəl</i>	<i>ddukkəl</i>	<i>ttdukkul</i>	“walk together; accompany”
CMB	<i>ddukk^wəl</i>	<i>ddukk^wəl</i>	<i>ttəddukk^wul</i>	“be friendly with”
Kabyle	<i>ddukəl</i>	<i>ddukəl</i>	[<i>tsdukəl</i>]	“walk together”

əcəcucəc*{20} əkərur(r)əy “to be round” [K: 750]**

Proto-Berber	Aorist <i>*əkərurəy</i>	Perfective <i>*ākərarāy</i>	Imperfective <i>*əttəkəruruy</i>	
Tuareg	<i>əkərurəy</i>	<i>[ākərarāy]</i>	<i>tikruruy</i>	“to be spherical”
Tashlhiyt	<i>krurri</i>	<i>krurri</i>	<i>[ttkrurri, ttkrurray]</i>	“to be round”

While this verb type is well-attested, both in varieties that have lost the short vowel distinctions, and those that have kept it, this verb is one of the few verbs of this type that is attested in both types of Berber languages, and can probably be reconstructed for Proto-Berber. Note, that there is a difference in length of the third root consonant *r* between Tuareg and Tashlhiyt.

{21} *əkənnunəy “to roll, tumble” [K: 750; B: 108]

Proto-Berber	Aorist <i>*əkənnunəy</i>	Perfective <i>*ākənnannāy</i>	Imperfective <i>*əttəkənnunuy</i>	
Ghadames	<i>gnunni</i>	<i>[āgənnune]</i>	<i>əttəgnunni</i>	“to tumble”
Tashlhiyt	<i>gnunni</i>	<i>[gnunni]</i>	<i>ttgnunnuy</i>	“to roll; tumble”
CMB	<i>çnunnəy</i>	<i>[çnunnəy]</i>	<i>ttəçnunnuy</i>	“to roll; tumble”
Mzab	<i>gnunni</i>	<i>[gnunni]</i>	<i>ttəgnunnuy</i>	“to roll on the ground”
Senhaja	<i>qnunni</i>	<i>[qnunni]</i>	<i>[ttəqnunni]</i>	“to roll”
Figuig	<i>qnunəy</i>	<i>[qnunəy]</i>	<i>ttəqnunuy</i>	“to roll”
Tarifit	<i>qnunni</i>	<i>[qnunni]</i>	<i>ttəqnunnuy</i>	“to tumble”

This semantically closely related verb is another verb that can be reconstructed for Proto-Berber. A variety of languages have expressive formations using the non-Proto-Berber phoneme *q* (Kossmann 2013b: 199ff.). While most varieties have a *g* as the first root consonant, there is no obvious explanation for the *k* in Central Moroccan Berber; whereas the *g* can be easily explained as the result of voicing assimilation, as such a reconstruction **əkənnunəy* is to be preferred over **əgənnunəy*.

{22} *əfərurəy “to drop fruits or grain (said of trees)” [B: 104; NZ: FRY1]

Proto-Berber	Aorist <i>*əfərurəy</i>	Perfective <i>*āfərarāy</i>	Imperfective <i>*əttəfəruruy</i>	
CMB	<i>frury</i>	<i>[frury]</i>	<i>tfruruy</i>	“drop fruits”
Tashlhiyt	<i>fruri</i>	<i>[fruri]</i>	<i>ttfruruy</i>	“id.”
Kabyle	<i>fruri</i>	<i>[fruri]</i>	<i>tstṣfruruy, [tstṣfruri]</i>	“id.”
Ouargla	<i>fruri</i>	<i>[fruri]</i>	<i>ttfruruy</i>	“to fall apart; decay”

{23} **əɣəlɣəl* “to be indecisive” [B 105; NZ: GL19]

Proto-Berber	Aorist <i>*əGəluGəl</i>	Perfective <i>*ǎGəlaGāl</i>	Imperfective <i>*əttəGəluGul</i>	
Tuareg	<i>gəlɣəl</i>	<i>əɣlagəl</i>	<i>tiglɣul</i>	“to be indecisive”
Ida U Semlal ¹⁵	<i>gluɣl</i>	<i>[gluɣl]</i>	<i>ttɣlɣul</i>	“to stagnate”

***əCəCucu**{24} **əɣəlulu* “to be hung”

Proto-Berber	Aorist <i>*əɣəlulu</i>	Perfective <i>*ǎɣələla</i>	Imperfective <i>*əttəɣəlulu</i>	
Ghadames	<i>əɣʷlulu</i>	<i>[ǎɣʷələla]</i>	<i>əttəɣʷlulu</i>	“to be hung”
Figuig	<i>ylulu</i>	<i>[ylulu]</i>	<i>ttəylulu</i>	“to be hung”

This verb is likely connected to the light verb **aɣəl* “to hang (s.th.)” (cf. Figuig *ayəl* “id.,” CMB *aɣəl* “id.,” Zng. P *yugäy*), but to my knowledge there are no other transitive/intransitive pairs that correspond to $ac^1əc^2$ and $əc^1əc^2uc^2u$ verb types. It is, therefore, difficult to decide how their derivational relationship should be understood.

***əC₁əC₂C₁əC₂**

It is quite difficult to find clear cognates across dialects that retain central vowel contrast and those that do not. However, it is clear that a marginally productive pattern of expressive reduplicative verbs exists and is productive across the different Berber varieties. And when these occur, they clearly have a $\Theta \rightarrow$ pattern in the **Aorist**, accompanied by an *i* vowel in the imperfective, as expected.

{25} **əffərɣər* “to flap the wings” [B: 172; NZ: FR1]

Proto-Berber	Aorist <i>*əffərɣər</i>	Perfective <i>*ǎffərɣār</i>	Imperfective <i>*əttəfərɣir</i>	
Ghadames	<i>əffərɣər</i>	<i>ǎffərɣār</i>	<i>əttəfərɣər</i>	“to flap the wings”
Kabyle	<i>ffərɣər</i>	<i>ffərɣər</i>	<i>ɪsfərɣir</i>	“id.”
Central Moroccan Berber	<i>fərɣər</i>	<i>fərɣər</i>	<i>tfərɣir</i>	“id.”
Tashlhiyt	<i>ffirɣir</i>	<i>ffirɣir</i>	<i>ttfirɣir</i> , <i>[ttfirɣar]</i>	“id.”
Figuig	<i>ffərɣər</i>	<i>ffərɣər</i>	<i>[tfərɣər]</i>	“id.”
Ouargla	<i>ffərɣər</i>	<i>ffərɣər</i>	<i>ttfərɣir</i>	“to fly around”

¹⁵ Basset (1929: 91), Boumalk (2003: 65, 172).

{26} **abb̥ar̥b̥ar* “to go out, travel” [NZ: BR2]

Proto-Berber	Aorist <i>*abb̥ar̥b̥ar</i>	Perfective <i>*ābb̥ar̥b̥ār</i>	Imperfective <i>*att̥ab̥ar̥bir</i>	
Tuareg (BF-M)	<i>ābb̥ār̥b̥ār</i>	<i>ābb̥ar̥b̥ār</i>	<i>att̥ab̥ār̥bar</i>	“to go out; leave”
Ouargla	<i>bb̥ar̥b̥ar</i>	<i>bb̥ar̥b̥ar</i>	<i>tt̥ab̥ar̥bir</i>	“to go up; travel”
Siwi	<i>bb̥ar̥baṛ</i>	<i>bb̥ar̥baṛ</i>	<i>[b̥ar̥baṛ]</i>	“to gush out”

Cases of Proto-Berber **b* are extremely rare, which is cause for Maarten Kossmann¹⁶ to doubt its reconstructibility. I agree with this assessment and, therefore, this verb and {27} **abb̥az̥b̥az̥* should not be uncritically accepted as reconstructible for Proto-Berber. However, both of these verbs are reduplicate and clearly expressive, onomatopoeic words. Onomatopoeia tends to allow for a broader range of phonemes than regular words (Dingemanse 2012), e.g., Japanese has a phoneme *p* in onomatopoeia and ideophones, while it is an allophone of *h* in native Japanese and Sino-Japanese words (McCawley 1968:77–80). Therefore, it can be imagined that verbs like these are indeed reconstructible with **b* despite the otherwise rarity of this phoneme.

{27} **abb̥az̥b̥az̥* “to shake, shiver”

Proto-Berber	Aorist <i>*abb̥az̥b̥az̥</i>	Perfective <i>*ābb̥az̥b̥āz̥</i>	Imperfective <i>*att̥ab̥az̥biz̥</i>	
Tuareg (M)	<i>ābb̥āḥ̣b̥āḥ̣</i>	<i>ābb̥aḥ̣b̥āḥ̣</i>	<i>att̥ab̥āḥ̣b̥aḥ̣</i>	“to shake; knock”
CMB (Ayt Atta)	<i>bb̥az̥β̣az̥</i>	<i>bb̥az̥β̣az̥</i>	<i>tt̥abb̥az̥β̣iz̥</i>	“to tingle”

The reconstruction of Proto-Berber **b* poses some problems (see {26} **abb̥ar̥b̥ar*, for a discussion).

{28} **add̥agd̥ag* “to shatter/to be shattered” [NZ: DG19]

Proto-Berber	Aorist <i>*add̥agd̥ag</i>	Perfective <i>*ādd̥agd̥āḥ</i>	Imperfective <i>*att̥add̥agd̥ig</i>	
Tuareg (M)	<i>[ādd̥āgd̥āg]</i>	<i>ādd̥agd̥āg</i>	<i>[att̥ādd̥āgd̥ag]</i>	“to shatter”
CMB	<i>dd̥aj̥d̥aj̥</i>	<i>dd̥aj̥d̥aj̥</i>	<i>tt̥d̥aj̥d̥ij̥</i>	“id.”
Kabyle	<i>dd̥aj̥d̥aj̥</i>	<i>dd̥aj̥d̥aj̥</i>	<i>add̥aj̥d̥ij̥</i> (< <i>*tt̥d̥agd̥ig</i>)	“id.”
Tashlhiyt	<i>dgd̥g</i>			“to be tired”
Ouargla	<i>dd̥agd̥ag</i>	<i>dd̥agd̥ag</i>	<i>tt̥d̥agd̥ig</i>	“to shatter”

The semantics of the Tashlhiyt verb are quite far removed from the rest but may be compared to the English expression “to be shattered” in the meaning of “to be

¹⁶Personal communication. Kossmann (1999: 128–130) suggests that several words appear to point to Proto-Berber **b*, but these are quite rare.

very tired” and Dutch “kapot zijn” (= to be broken) – in the meaning of “to be very tired.”

This verb is widely attested in North-African Arabic, e.g., Moroccan *dəgdəg* “to shatter into tiny pieces,” *mdəgdəg* “shattered; very tired” (de Prémare et al. 1993–1999: IV, 302), Tunisian *mdəgdig* “broken; shattered”.¹⁷ It is difficult to decide the direction of borrowing, or whether this is a case of parallel development of an onomatopoeic word.

{29} **attəftəf* “to grope around”

Proto-Berber	Aorist <i>*attəftəf</i>	Perfective <i>*əttəftəf</i>	Imperfective <i>*əttəftəf</i>	
CMB	<i>əftəf</i>	<i>əftəf</i>	<i>əftəf</i>	“to grope around”
Kabyle	<i>ttəftəf</i>	<i>ttəftəf</i>	<i>ttəftəf</i>	“id.”
Ouargla	<i>ttəftəf</i>	<i>ttəftəf</i>	<i>ttəftəf</i>	“id.”

**əcəcicəc*

{30} **əgəriwəl* “to turn” [B: 96; NZ: GRWL]

Proto-Berber	Aorist <i>*əgəriwəl</i>	Perfective <i>*əgəriwəl</i>	Imperfective <i>*əttəgəriwil</i>	
Tuareg	<i>əgəriwəl</i>	<i>əgəriwəl</i>	<i>əttəgəriwil</i>	“to turn around”
Tashlhiyt (Ntifa) ¹⁸	<i>griwl</i> , [grawl]	<i>grawl</i>	<i>ttgriwil</i>	“id.”

Already at the Proto-Berber stage **uw* had dissimilated to *iw* (Prasse 1973: 93), therefore this verb and {31} **əfəriwəs* may have originally belonged to the **əcəcicəc* verbs, discussed above.

{31} **əfəriwəs* “jump (with joy?)” [FRWS1,2]

Proto-Berber	Aorist <i>*əfəriwəs</i>	Perfective <i>*əfəriwəs</i>	Imperfective <i>*əttəfəriwis</i>	
Tuareg	<i>əfəriwəs</i>	<i>əfəriwəs</i>	<i>əttəfəriwis</i>	“to be joyful”
Kabyle	<i>firiwəs</i>	<i>firiwəs</i>	<i>ttəfəriwis</i>	“to jump up”

The semantic connection of the Kabyle verb “to jump up” and Tuareg “to be joyful” is not perfect but seem close enough for them to be cognates. In any case, the formal similarity is perfect, allowing us to examine the formation of a **əcəcicəc* verb.

¹⁷ From the TUNICO dictionary (<https://vicav.acdh.oeaw.ac.at/>) (consulted 8 january 2020).

¹⁸ Basset (1929: 86, 220).

{32} **arə́ǵiǵəʔ* “to tremble” [K: 323; B: 97]

Proto-Berber	Aorist <i>*arə́ǵiǵəʔ</i>	Perfective <i>*ä́rə́ǵaǵaʔ</i>	Imperfective <i>*ə́ttarə́ǵiǵiʔ</i>	
Tuareg	<i>ərgiǵi</i>	<i>ərglagla</i>	<i>ttirgiǵi</i>	“to tremble”
Figuig	<i>rziʒ</i>	<i>[rziʒ]</i>	<i>ttərziʒi</i>	“id.”
Ouargla	<i>rziʒi</i>	<i>[rziʒi]</i>	<i>ttərziʒi</i>	“id.”
Zenaga	<i>ərgəgi</i>	<i>ärgägä</i>	<i>ttərgəgi</i>	“id.”
Awjila	<i>ərgig</i>		<i>[tərgiga]</i>	“id.”
Kabyle	<i>rjiji</i>	<i>rjaji</i>	<i>ts̥ts̥ərjiji</i>	“id.”
CMB	<i>rjiji</i> <i>rjiggi</i>	<i>[rjiji]</i> <i>rjigga</i>	<i>ttərjiji</i> <i>ttərjiggi</i>	“id.”
Tashlhiyt	<i>rgigi</i>	<i>rgagi</i>	<i>ttrgigi</i> <i>ttrgigiγ</i>	“id.”

Kabyle, Central Moroccan Berber and Tashelhiyt have reinterpreted this verb as having a final root consonant **y*: **arə́ǵiǵəy*, as is also evidenced by the verbal noun Kb. *θarjajayθ*, Tashl. *targigayt* (but CMB *targagit*). The A→ vocalism attested in Zenaga suggests that perhaps Zenaga underwent a syncope of the initial open syllable, similar to many Tuareg dialects with verbs of this type (e.g., Tu. *kəniḥər* but P *āknaḥər* < **ākənaḥər* “be disgusting”). The Figuig and Awjili reflexes rather support a reconstruction with **v*ʔ, as final **i* or **əy* would be expected to yield *i* or *əy*.

{33} **əmə́liləʔ* “to meet one another”

Proto-Berber	Aorist <i>*əmə́liləʔ</i>	Perfective <i>*ä́mə́laläʔ</i>	Imperfective <i>*ə́ttəmə́liliʔ</i>	
Kabyle	<i>mlil</i>	<i>mlal</i>	<i>ts̥ts̥əmlili</i> , <i>[ts̥ts̥əmlil]</i>	“to meet”
CMB (Ayt Atta)	<i>lmili</i>	<i>lmala</i>	<i>ttəlmili</i>	“to be close by”

The Ayt Atta form is likely a metathesis of this root, rather than the Kabyle form, as **ə́c₁ə́c₂ic₂əʔ* reduplication seems common in verbs of this shape (see also {32} above).

Bibliography

- Basset A (1929) La langue berbère. Morphologie, le verbe, étude de thèmes. Librairie Ernest Leroux, Paris
- Benamara H (2013) Dictionnaire amazighe – français: parler de Figuig et ses régions. Institut royal de la culture amazighe, Rabat
- Boumalk A (2003) Manuel de conjugaison du tachelhit. L'Harmattan, Paris
- Cohen D, Taine-Cheikh C (2000) À propos du zénaga. Vocalisme et morphologie verbale en berbère. Bulletin de la Société de Linguistique de Paris 95:267–319
- Dallet J-M (1953) Le Verbe Kabyle. Fichier de Documentation Berbère, Fort National
- Dallet J-M (1982) Dictionnaire kabyle-français. parler des At Mangellat, Algérie. SELAF, Paris
- de Foucauld C (1950) Dictionnaire touareg-français (dialecte de l'Ahaggar). Imprimerie Nationale de France, Paris

- de Prémare A-L, Colin GS, Brunot L, Marçais W (1993–1999) Dictionnaire Arabe-Français. Établi Sur La Base de Fichiers, Ouvrages, Enquêtes, Manuscrits Études et Documents Divers Par A.-L. de Prémare. 12 vols. L'Harmattan, Paris.
- Dell F, Elmedlaoui M (1991) Clitic Ordering, Morphology and Phonology in the Verbal Complex of Imdlawn Tashlhiyt Berber. Part II. Langues orientales anciennes philologie et linguistique. Peeters, Louven and Paris
- Dingemanse M (2012) Advances in the cross-linguistic study of ideophones. *Lang Linguist Compass* 6(10):654–672
- Heath J (2005) A grammar of Tamashek (Tuareg of Mali). Mouton de Gruyter, Berlin
- Heath J (2006) Dictionnaire touareg du Mali. tamachek-anglais-français. Karthala, Paris
- Ibañez E (1959) Diccionario Español-Senhayi (Dialecto Bereber de Senhaya de Serair). Instiuto de Estudios Africanos, Madrid
- Kossmann MG (1997) Grammaire du parler berbère de figuig. Peeters, Paris/Louvain
- Kossmann MG (1999) Essai sur la phonologie du proto-berbère. Rüdiger Köppe Verlag, Köln
- Kossmann MG (2000) La futur à Ghadamès et l'origine de la conjugaison verbale en berbère. In: Chaker S, Zaborski A (eds) Études Berbères et chamito-sémitiques. mélanges offerts à Karl-G. Prasse. Peeters, Paris/Louvain, pp 237–255
- Kossmann MG (2001a) The origin of the glottal stop in Zenaga and its reflexes in the other Berber languages. *Afrika und Übersee* 84:61–100
- Kossmann MG (2001b) L'origine du vocalisme en zénaga de Mauritanie. *Frankfurter Afrikanistische Blätter* 13:83–95
- Kossmann MG (2008) Three irregular berber verbs: 'eat', 'drink', 'be cooked, ripen. In: Lubotsky A, Schaeken J, Wiedenhof J (eds) Evidence and counter-evidence, Festschrift Frederik Kortlandt. Rodopi, New York, pp 225–236
- Kossmann MG (2011) A grammar of Ayer Tuareg (Niger). Rüdiger Köppe Verlag, Köln
- Kossmann MG (2012) Berber. In: Frajzyngier Z, Shay E (eds) The Afroasiatic languages. Cambridge University Press, Cambridge, pp 18–101
- Kossmann MG (2013a) A grammatical sketch of Ghadames Berber (Libya), *Berber studies* 40. Rüdiger Köppe, Köln
- Kossmann MG (2013b) The Arabic influence on northern Berber. Brill, Leiden/Boston
- Kossmann MG (2017) La Place Du Parler Des Senhaja de Sraïr Dans La Dialectologie Berbère. In: Vicente Á, Caubet D, Naciri-Azzouz A (eds) La Région Du Nord-Ouest Marocain. Parlers et Pratiques Sociales et Culturelles. Prensas de la Universidad de Zaragoza, Zaragoza, pp 93–108
- Kossmann MG (2018) Zenaga reflexes of Berber final weak verbs. *Folio Orientalia* 55:161–194
- Kossmann MG, Suchard DB (2018) A reconstruction of the system of verb aspects in proto-Berber-Semitic. *Bull Sch Orient Afr Stud* 81(1):41–56
- Lafkioui M (2018) The imperfective in Berber. In: Tosco M (ed) Afroasiatic. Data and perspectives. John Benjamins, Amsterdam, pp 85–103
- Lanfry J (1973) Ghadamès. étude linguistique et ethnographique. II. Glossaire (parler des Ayt Waziten). Fort-National, Algérie
- Mauri S (2015) Formal and functional properties of grammatical aspect in Ayt Atta Tamazight. PhD Thesis, The School of Oriental and African Studies, London.
- McCawley JD (1968) The phonological component of a grammar of Japanese. Mouton, The Hague
- Mourigh K (2016) A grammar of Ghomara Berber (North-West Morocco), *Berber studies* 45. Rüdiger Köppe, Köln
- Naït-Zerrad K (1998) Dictionnaire Des Racines Berbères (Formes Attestées) I. A-BEẒL. Paris/Louvain, Éditions Peeters
- Naït-Zerrad K (1999) Dictionnaire Des Racines Berbères (Formes Attestées) II. C-DĒN. Paris/Louvain, Éditions Peeters
- Naït-Zerrad K (2002) Dictionnaire Des Racines Berbères (Formes Attestées) III. D-GEY. Paris/Louvain, Éditions Peeters
- Oussikoum B (2013) Dictionnaire amzighe – français: le parler des Ayt Wirra, Moyen Atlas, Maroc. IRCAM, Rabat

- Prasse K-G (1973) Manuel de grammaire touarègue (tāhaggart) 2: VI–VII: Verbe. Éditions de l'Université de Copenhague, København
- Prasse K-G, Alojaly G, Mohamed G (1998) Lexique Touareg – Français. Deuxième édition revue et augmentée. Museum Tusculanum Press, København
- Serhoual M (2002) Dictionnaire Tarifit-Français. PhD Thesis, Université Abdelmalek Essaâdi, Tétouan.
- Souag L (2017) La Diffusion En Berbère : Reconcilier Les Modèles. In: Diffusion, implantation, convergence, Mémoires de La Société de Linguistique de Paris, N.S. 24. Peeters, Louvain, pp 83–107
- Stroomer H (Forthcoming) Dictionnaire tašlhiyt-français
- Sudlow, D. (2001). The Tamasheq of North-East Burkina Faso. notes on grammar and syntax including a key vocabulary. Köln: Rüdiger Köppe Verlag.
- Sudlow D (2009) Dictionary of the Tamasheq of north-East Burkina Faso. Rüdiger Köppe Verlag, Köln
- Taifi M (1992) Dictionnaire tamazight-français (parlers du Maroc central). L'Harmattan-Awal, Paris
- Taine-Cheikh C (1999) Le Zénaga de Mauritanie à La Lumière Du Berbère Commun. In: Lamberti M, Tonelli L (eds) Afroasiatica Tergestina. Paper from the 9th Italian Meeting of Afro-Asiatic (Hamito-Semitic) Linguistics, Trieste April 23–24, 1998. Unipress, Padova, pp 299–324
- Taine-Cheikh C (2008) Dictionnaire zénaga-fançais. Rüdiger Köppe Verlag, Köln
- Taine-Cheikh C (2010) Dictionnaire français-zénaga (berbère de Mauritanie). Rüdiger Köppe Verlag, Köln
- Van den Boogert N, Kossmann M (1997) Les premiers emprunts arabes en berbère. Arabica 44: 317–322
- van Putten M (2014) A grammar of Awjila (Libya): based on Paradisi's work. Rüdiger Köppe Verlag, Köln
- van Putten M (2018) Proto-Berber mid vowel harmony. Nord J Afr Stud 27(3):1–32

Vermondo Brugnatelli

Contents

Evidence of <i>isnin</i> and Other Defined Numerals	337
Beni Menacer	337
Aurès	337
Kabyle	339
Tripolitania	341
Jerba	341
Mزاب	342
Touareg	343
Southern and Northern Morocco: <i>Tashelhit</i> , <i>Iznassen</i>	346
Zenaga	347
The <i>Kitāb al-Barbariyya</i>	349
Diachronic Considerations	351
Conclusions	354
References	355

Numerals, along with other quantifiers such as *many* or *few*, are generally assigned to the word class of indefinite determiners (example 1.a). However, due to their peculiar nature, they also share some features of adjectives and nouns and may, consequently, receive marks of definiteness, like definite articles in languages that possess them (examples 1.b and 1.c). On the peculiar nature of numerals, see, among others, Brugnatelli (1982: 121–125) and the literature cited there. As far as Berber numerals are concerned, Galand (1967: 253) emphasizes their nominal nature: “Le nom de nombre en berbère est bien un nom, pourvu d’une morphologie particulière mais capable d’assumer à lui seul une ‘ fonction primaire.’”

V. Brugnatelli (✉)
 Università di Milano-Bicocca, Milan, Italy
 e-mail: vermondo.bugnatelli@unimib.it

- (1) a. Indefinite phrases:
many people
few people
three people
- b. Definite phrases (“strong”)
the three people (already mentioned)
the three (already mentioned)
- c. Definite phrases (“weak”)
the three of us (three people being involved in a conversation)

This paper aims at analyzing the behaviour of Berber numerals with regard to both kinds of “definiteness” singled out by recent research: anaphoric (or “strong”) and unique (or “weak”). Anaphora is a well-known and studied concept: a means of “cohesion,” linking an utterance to something already mentioned (or alluded to) in the discourse. The phrase *two people entered France with a tourist visa: both pretend to be farmers* shows a typical case of numerals in anaphora: *both* = “the two already mentioned.” On the other hand, the so-called unique definites, “pick up a referent which is the single or maximal element in the property denoted by their restrictor” (Cheng et al. 2017: 79), and this is typically the case for numerals accompanied by a personal pronoun: defined despite the lack of an earlier mention of the number word, the expression *the three of us* is defined, since three people are concerned and the utterance embraces the totality of them (in such cases all may replace or accompany the article: *all three of us*).

Unlike English and many other European languages which have definite and indefinite articles, the Berber languages lack articles and do not systematically express whether a noun phrase is definite or indefinite. When it comes to stressing the definiteness in a sentence, these languages use specific means, above all demonstratives. Nevertheless, peculiar properties concerning numerals in both cases of definiteness appear in several Berber languages. Some of them use special forms of numerals in anaphora with reference to a known number of items. Moreover, numerals followed by personal pronouns, expressing the totality of involved items, often require the insertion of a special segment restricted to these constructions.

More specifically, several languages possess a word meaning “both” that appears most often as *isnin* and is undoubtedly related to the numeral *sin* (or *sən*) “two.” This expression was first singled out by André Basset in a meeting of the *Groupe Linguistique d’Études Chamito-Sémitiques* (GLECS), held in Paris on March 27, 1946. He detected this term in the Berber language of the Aurès mountains (also known as Shawia) and described some of its properties. His study also features a comparison with analogous expressions found in Tuareg and Kabyle, aimed to explain the origin of this word. (A. Basset (1945–1948; see also 1969: 39).)

In his paper, Basset treats *isnin* as a unique feature, as it seemed only attested in Shawia Ait Fraḥ, but it turns out that this term also appears elsewhere: several other languages share this expression. Therefore, it can be assumed that it belongs to an ancient phase of the Berber languages. In this paper, I intend to resume the survey

started by Basset and subsequently neglected by Berberologists. Considering the new data available, I will try to investigate the origin of this expression.

Evidence of *isnin* and Other Defined Numerals

As far as I know, the presence of *isnin* or similar expressions appears in languages and dialects from Algeria (Beni Menacer, Aurès and Harakta, Kabylia, Mzab, Northern Tuareg), Tunisia (Jerba), Libya (Zuara and Tripolitania), Mali, and Burkina Faso (Southern Tuareg). In addition to this contemporary evidence, *isnin* is also attested in the medieval language of the *Kitāb al-Barbariyya* (an Eastern dialect between Libya and Tunisia).

Beni Menacer

An old record, possibly the earliest, of a specific term for “both,” noted *snin*, was taken among the Beni Menacer (near Cherrhell, Algeria) and published within the studies that René Basset, André’s father, devoted to their language in the nineteenth century (2):

- (2) oud’ern d’is snin ard’akhel

وذن ديس سنين ار داخل

“they both went down inside” (R. Basset 1884: 207–8)

rouh’an r’ elkhela d’i senin

روحن غالخلا ذي سنين

“they both went to the fields”

iâredh iouarrachen enni d’i senin

يعرض يواراشن اني ذي سنين

“he met those children both of them” (R. Basset 1892b: 49, 52). (The original French phrases are “ils descendirent tous deux à l’intérieur,” “ils allèrent aux champs tous deux,” and “il rencontra les enfants ceux-là tous deux.”)

The notations by Basset do not rule out the possibility that the word was actually *isnin*, whose initial *i-* was typographically separated and attached to the preceding particle *d* (or *ḍi*) by the influence of the numeral “two,” which normally is recorded as *sen* (once *sin*), f. *senath* and *senat* (probably [sən], [sənaθ]/[sənat]).

Aurès

The attestations of this term in the Aurès region are also quite old. The earliest records seem to be those taken by René Basset among the Harakta in the late nineteenth century. In his collection of tales entitled *Loqman berbère*, one finds in fact:

(3) *k'imen essaoullen f elk'aouth ensen snin*

قيمن اساوّلن فالقوث نسن سنين

“they both started talking about their strength” (R. Basset 1890: 48). (“Ils se mirent à parler de leur force à eux deux.”)

According to René Basset, the text was elicited from Ali ben H'amed (of Oulad Sidi Younes) in 1887. The item *snin* also appears, both in the glossary at the end of the book (R. Basset 1890: 271), and in that of a short paper he wrote on the Harakta (1892a: 9).

This term also occurs within the Shawia texts collected and published by G. Mercier in 1900 in Ahmar Khaddou, recently verified and transcribed by M. Lafkioui and D. Merolla (2002).

(4) *ir'ersasen issenin* “he slaughtered them both” (Mercier 1900: 211, 214)

iyers-asen i sn-in (Lafkioui and Merolla 2002: 58–59)

h'elan issenni “both being beautiful” (Mercier 1900: 218, 226)

hlan i-s senni (Lafkioui and Merolla 2002: 62–63)

eidhoun issenin “they both fell” (Mercier 1900: 242, 248)

yqun is snin (Lafkioui and Merolla 2002: 90–91). (The original French phrases are “il les tua tous les deux,” “étant beaux tous deux,” and “ils tombent tous deux.”)

Mercier remarked in a note: “*sen* “two,” *issenni* “both,” is a determined plural noun,” (Mercier (1900: 218): “*sen* ‘deux’, *issenni* ‘tous les deux,’ est un substantif pluriel déterminé.”) referring to the form ending in *-i*, while he did not comment on the forms ending in *-in* of the other occurrences.

The most extensive studies on Shawia to date are those that André Basset devoted to the dialect of the Ait Frah. His findings in this Berber language prompted him to draw the attention of Berberologists to these special numerals, m. *isnin*, f. *tisnin* “tous les deux” (A. Basset 1945–1948: 19). Later, in his collection of texts, published posthumously in 1961, he records a longer series of defined numerals, from “two” to “ten,” both in the masculine and in the feminine (A. Basset 1961: 127): (This list lacks the numeral corresponding to ‘all eight’, possibly due to an oversight. For the sake of comparison, I have joined a column with the corresponding standard numerals (A. Basset 1961: 7).)

(5)	Masculine	Feminine	Standard numerals
“all two”	<i>isnin</i>	<i>tisnin</i>	m. <i>sən</i> f. <i>sənt</i>
“all three”	<i>itlataggən</i>	<i>tītlataggin</i>	c. <i>tlatā</i>
“all four”	<i>irḥəaggən</i>	<i>tirḥəaggin</i>	<i>rḥəa</i>
“all five”	<i>iḥəmsəggən</i>	<i>tiḥəmsəggin</i>	<i>ḥəmsa</i>
“all six”	<i>istəggən</i>	<i>tistəggin</i>	<i>sətta</i>
“all seven”	<i>isḥəaggən</i>	<i>tiḥəaggin</i>	<i>səbea</i>

“all nine”	<i>itsaεaggən</i>	<i>ʔitsaεaggin</i>	<i>ts ʔa</i>
“all ten”	<i>iešərəggən</i>	<i>ʔiešərəggin</i>	<i>ε ʃra</i>

These elements were also described and analyzed by Th. Penchoen in his structural systematization of the materials collected by A. Basset, but in his description, he only reports the forms of “both” and “all three” (Penchoen 1973: 149 § 10.24), the only appearing in context. In (6), one can see three examples taken from Basset’s texts, the last two reported also in Penchoen’s analysis:

- (6) *bɔun ɖəssən isnin* (A. Basset 1961: 333, l. 49)
 “they both burst laughing”
təɖərn-ihənt tɪsnin (A. Basset 1961: 25, l. 23)
 “she turns both of them”
grin-hən itlatəggən dəg-giğgən-ttuggʷ (A. Basset 1961: 124, l. 20)
 “they put down all three of them at once”

Besides, a quick survey I made on the Internet among Berber speakers from the Aurès region revealed some other interesting features not yet recorded in the existing literature, as far as I know:

1. For the feminine of *isnin*, besides *tɪsnin/tɪsnin*, there also exists a form *isntin* (several informants agree on this).
2. The forms noted by Basset for the numerals “three” to “ten” are still living: “all 3, 4, 5 ... we say: *itlaytiyen*, *irabɛiyen*, *ixamsiyen* ... these are berberised borrowings” (informant J.O. 20-1-2012). (“Tous les 3, 4, 5 ... on dit: *itlaytiyen*, *irabɛiyen*, *ixamsiyen*. ... ce sont des emprunts berbérésés.” The morpheme *-əggən* recorded by Basset displays a phonetic realization *-gg-* of geminated *y* (so it corresponds to **-əyyən*).

Interestingly, a series of defined numerals from “two” to “ten” can be observed in Mزاب as well and will be dealt with below.

Kabyle

The presence of *isnin* in Kabylia (Northern Algeria) has been pointed out by several authors, in different areas of this region. For this reason, Prasse expresses his surprise at the lack of this lexeme in the huge dictionary of J.-M. Dallet (1982): “Curiously, some other words also are not found in the dictionary, for instance (...) *isnin* ‘both’ (= *i-sin*). *isnin* is probably not used in the variety of At-Mangellat, but it is known in other Kabyle dialects and is so essential in the Berber context in general that it should have been incorporated” (Prasse 1984: 170). (“Quelques autres mots aussi, curieusement, ne se trouvent pas dans le dictionnaire, p. ex. *ɾwah!* ‘viens,’ *isnin* ‘tous les deux’ (= *i-sin*). *isnin* probablement n’est pas usité dans le parler des At-Mangellat, mais il est connu dans d’autres parlers kabyles et est tellement

essentiel dans le contexte berbère en général qu'on aurait dû l'incorporer.”) In the following example (7), I report some instances from the western dialect of Ait Jennad (7.a) and the eastern dialects of Aokas (7.b) and Ait Abbas (7.c). According to several informants, it seems that nowadays *isnin* is still used in the Eastern part of Kabylia only, while it completely disappeared in the Western part.

- (7) a. *Mer d' emmis tejed'aount, thili isk'arni imezzour'en is i senin. Armi entsa ahayin, iouen oumezzour' iuod'er [= mār d əmmis n təjəðeunt, tili yəsqarni iməzzuyən-is isnin. Armi natta a-h-a-yin, yiwən uməzzuy yuḍər]*
 “If he were the son of the mare, then he would prick both ears. While as to him, one ear is down” (Mouliéras 1999: 110–111).
- b. *Ma izra-t niy iṭṭəf-in s isnin, d'in-iy*
 “If he sees him (his wife's lover) or catches them both, he kills them” (Berkai 2012–2013: 903).
- c. *ylin isnin d əlməyytin*
 “They both dropped dead” (At-Abbas 1976: 65).

An ancient occurrence of *isnin* also appears in a religious text recovered from a manuscript studied by E. Gutova and dating back to the eighteenth or nineteenth century: *Iğərm ššəffaf, u ṭn-imṃi ara isnin: ur imṃi akəččum, ur imṃi asəkkud* “transparent [and not solid] matter does not stop either (of these): it does not stop entering and it does not stop seeing through” (lines 1.7–8 of the manuscript). In her linguistic analysis of the manuscript, she puts forward the hypothesis that this form is “a blend of Berber *sin* and Arabic *iṭnain*” (Gutova 2011: 41, 132). (Properly speaking, this is a cataphora. Gutova (2016) contains a partial publication of this text; a full publication is in preparation.)

A feminine form is hard to find. The only record I am aware of is *tisnin*, reported by Prasse for comparison in his grammar of Tuareg (1974: 408), and identical to the Aurès numeral studied by Basset. (In accordance with the principle “*testis unus, testis nullus*”, I am tempted to regard Prasse's attribution of *tisnin* to Kabyle as a slip of the pen.) Nowadays, it seems that the only form attested in Kabyle is the masculine one. I have found no trace of a feminine counterpart, either in texts or through informants. Whenever it is necessary to express “both” related to females or feminine nouns, only *snat* is normally used, following the usual way of expressing “all X.”

Interestingly, this construction, namely, *i* + numeral, occurs with defined numerals only and is distinctly different from the way of attributing a numerical quantity to a given item: *ruḥən i ṭlata* “all three left” versus *ruḥən di ṭlata* “they left, three (of them).” (“Ils sont partis tous les trois” vs. “ils sont partis au nombre de trois”) (Dallet 1982: 354, 133, respectively.) From a formal point of view, the morpheme *i* that characterizes this construction could safely be considered a prefix, rather than a separate preposition, given that nothing can be placed between it and the numeral. As a consequence, one could consider Kabyle too as a language possessing a category of “defined numerals” *isnin/isin* “both,” *iṭlata* “the three”/“all three,” *irəbea* “the four”/“all four,” etc. It would be quite similar to the series of the Aurès region, with the same prefix but devoid of the ending *-iyən/-əggin*.

Tripolitania

In the Berber language of Zuara, on the coast of Tripolitania (Libya), the existence of *isnin* is attested in the texts collected by Mitchell in the Fifties and published shortly after his death: *aləm̄mi tər̄lim isnin xir* ‘it is better for both of you to be content’ (Mitchell 2009: 284).

Some decades earlier, Beguinot already noted this word in an isolated utterance within a paper on some conventional expressions in Tripolitanian Berber: *if̄assn-ik isnin* ‘ten’, literally ‘both your hands’ (Beguinot 1917–1918: 111).

A quick survey on the Internet has provided other occurrences, not only of the masculine form *isnin* (8.a), but also of its feminine counterpart *tisəntin* (8.b):

- (8) a. *utəfən i Fransa s lvižət siyaḥiyya ass n 20 ayustəs: A.F.B 31 n isəggasən, M. S. S. 29: yih isnin xsən iman nnsən d ifəllahən* “on August 20 entered France with a tourist visa: A.F.B. 31 years old and M.S.S., 29 years old: they both pretend to be farmers”
- b. *yisi dis iğən d əəiraqi d aməqqar g Drabləs, ittutlay did ləḥkumət n Libya af irra n lealaqat jar timura tisəntin* “there was a prominent Iraqi in Tripoli who was speaking with the Libyan government about opening relations between the two countries.” (Sentences taken from the blog www.facebook.com/1042773265869656/posts/1315950768551903/ (last visited 27 December 2020). Translation is mine, as well as a light editing of the Berber orthography. I could also ascertain the existence of *isnin* and *tisentin* from a Zuaran informant.)

In the rest of Berber Tripolitania (Jebel Nefusa), *isnin* has never been detected in the existing studies. Nevertheless, a researcher who has done much fieldwork in that region, and still is in touch with individuals and groups of local Berber speakers, Nora Mahi, kindly informed me that in El-Qalaa, next to Yefren, in the eastern margin of the Jebel, she observed a frequent use of similar expressions for masculine and feminine nouns as well. The masculine form is *isnin*, and the feminine is *isnintən*. Here is a sample of the use of these expressions in El-Qalaa:

- (9) *ufix-tən isnin g aylad* “I found them both in the street”
isnintən awwant-iyi-d awal “both (women) repeated the same speech to me”

Jerba

Similarly, the Tunisian language of Jerba too, quite close to the Libyan border, also records the masculine (10.a) and feminine (10.b) forms of a specific term for “both; the two,” identical to those recorded in Zuara.

The following examples have been recorded by André Basset in the Thirties. I have found them within some of his handwritten notes, kept in the BULAC library in Paris, and still unpublished:

- (10) a. *usən-d s isnin* “they (m.) came, both of them”
 b. *usənt-t sənt tsədnan tisəntin* “two women came, both of them.”
 (Compare the analogous expressions I have recorded from contemporary speakers: *usen-d isnin* and *usent-ed tisentin*.)

Mzab

Interestingly, definite numerals in *Tumzabt*, the language spoken in the Mzab region in Algeria, show peculiar aspects.

The numeral *isnin* does not appear as such in the dictionaries and, up to now, I have only found it in a folktale translated into *Tumzabt* from a Tunisian Berber text published by Stumme (1900): *qqimən ssusmən d isnin* “and they both fell silent” (Bekkay 2012: 21). (The corresponding Tunisian Berber passage was *sūsmeṇ sīn lsénn* ‘dann schwiegen die Beiden’ (Stumme 1900: 1, 43). The translation into *tumzabt* was made from a French translation (“et ils se turent tout les deux,” Bekkay 2012: 24).)

Some recent dictionaries, like Delheure’s (1984), record a variant of this term, m. *sənnin* f. *sənnətin*, in the following construction:

ay sənnin “both of them (m.)” (“tous les deux”)
ay sənnətin “both of them (f.)” (“toutes les deux”).

The comparison with the other Berber languages possessing a similar word shows the lack of the initial *i-* (or *ti-* for the feminine), which is replaced by an invariable particle *ay*. This particle, according to the same dictionary, is a “pronominal demonstrative element that can be used as a “support of determination” before a noun, a pronoun, a relative clause.” (In French, “support de détermination.” This terminology was coined by L. Galand (1969: 96) “J’appelle support de détermination un nominal dont la fonction propre est de recevoir un déterminant, élément démonstratif, proposition relative ou complément déterminatif: ainsi le français *celui* dans *celui-ci*, *celui que j’ai vu*, *celui de Jean*”. Some authors use the English expression “pronominal head.”) Consequently, the original meaning of *ay sənnətin* would be: “ce de deux = tous les deux, à deux”, according to Delheure (1984: 241).

A peculiar feature of *Tumzabt* is the existence of a full series of “defined numerals” beyond “two,” and up to “ten.” The same construction occurring with “two” is adopted for all these numerals: *ay* + numeral + the suffix *-in*.

(11)	Standard numerals (m. and f.)	defined numerals (m. and f.)
2	m. <i>sənn</i> , f. <i>sənnət</i>	<i>ay sənnin</i> , f. <i>ay sənnətin</i> “the two; both”
3	m. <i>šarəḍ</i> , f. <i>šarəḍt</i>	<i>ay šarḍin</i> , f. <i>ay šarəḍtin</i> “the three”
4	m. <i>əkk^wəz/ukk^wəz</i> f. <i>əkk^wəzt</i>	<i>ay əkk^wəzin</i> , f. <i>ay əkk^wəztin</i> “the four”
5	m. <i>səmməs</i> , f. <i>səmməst</i>	<i>ay səmməsin</i> , f. <i>ay səmməstin</i> “the five”
6	m. <i>səšš</i> , f. <i>səššət</i>	<i>ay səššin</i> , f. <i>ay səššətin</i> “the six”
7	m. <i>sa</i> , f. <i>sat</i>	<i>ay sayin</i> , f. <i>ay satin</i> “the seven”
8	m. <i>tam</i> , f. <i>tamət</i>	<i>ay tamin</i> , f. <i>ay tamtin</i> “the eight”
9	m. <i>təšš</i> (<i>təzz</i>), f. <i>təššət</i>	<i>ay təššin</i> , f. <i>ay təššətin</i> “the nine”
10	m. <i>mṛaw</i> , f. <i>mṛawt</i>	<i>ay mṛawin</i> , f. <i>ay mṛawtin</i> “the ten”

Example number (12) shows some utterances containing these defined numerals:

- (12) *yutəf-d s wəylad, yəščəwčəw-as i baba-s taməzzuyt-əs, fḥəyan ay snin*
 “he walked in from the street, whispered in his father’s ear and they both left.” (*Iwalen* n.d.: 28)
bdan ttmənyan, ay sənnin-u, ammas n wəzɣar uɣərm
 “these two started arguing (and fighting) in the middle of the marketplace” (*Iwalen* n.d.: 215)
a-isəntin təqqimanət f-idis n-ižəžən, iggət aməldad m-tidiqət
 “they (f.) both sit by the stakes, facing each other” (Aulard 1989: 145)

Touareg

In his quest for the origin of *isnin*, A. Basset turned to Tuareg, the only language where a special category of defined numerals had already been described by Foucauld (1920: 103–104). (Moreover, some of these terms appeared in context within the *Textes Touaregs en Prose* published in 1922. In his paper, Basset refers to the following passages (here quoted from Prasse’s revised transcript of 2010): *təqqəd-i təfuk dəy fassān-in issənān* ‘the sun burnt me on both hands’, text 44 (159) p. 24; *ywāt-t dəy elyan issənān* (TP *əssənān*) ‘he hit him on both legs’, text 147 (124) p. 86; *āmmūtān ikkārḍān (əkkārḍən?) āmōsān meddān* ‘all three died in adulthood’, text 136 (106) p. 76.) The term adopted by Prasse (1974: 408) for these numbers is “nombres collectifs” (collective numerals). Unlike the series observed in Tumzabt, the Tuareg numerals are a limited set, from “two” to “four,” possibly the remnants of a longer series (“débris d’une série”, Prasse *ibid.*). In Tahaggart (Tuareg language of southern Algeria), they are:

(13)	defined numerals	versus	standard numerals
	m. <i>əssənən</i> f. <i>əssənätīn</i>	“both”	m. <i>əssin</i> f. <i>sänât(ät)</i>
	m. <i>əkkərḍən</i> f. <i>əkkərḍätīn</i>	“all three”	m. <i>kāraḍ</i> f. <i>kārâḍät</i>
	m. <i>əkkəzən</i> f. <i>əkkəzätīn</i>	“all four” (In his description, Foucauld remarks that the feminine forms may occur, both with and without an initial <i>t</i> - (the typical morpheme of feminine in nouns): “ <i>essenetīn</i> , <i>ekkerḍetīn</i> , <i>ekkezetīn</i> and <i>tessenetīn</i> , <i>tekkerḍetīn</i> , <i>tekkezetīn</i> are equally correct, the former forms being more frequent than the latter” (Foucauld 1920: 104). The vocalism given by Foucauld might be inaccurate, as Prasse repeatedly pointed out. According to Heath (2005: 256), the forms for “both” presumably represent <i>əssən-än</i> and (<i>t</i>) <i>əssän-ät-en</i> . The most recent transcription of Foucauld’s <i>Textes Touaregs en Prose</i> by Prasse (2010) contains the forms <i>issənăn</i> (TP <i>əssənăn</i>) p. 24, 86 and <i>ikkārḍăn</i> (<i>əkkārḍən?</i>) p. 76.)	m. <i>ökkoz</i> f. <i>ökkôzät</i>

Similar series – always from “two” to “four” – are attested in Southern Tuareg too. (At least in Mali and Burkina Faso. I have not found these peculiar forms in the Tuareg dialects of Niger.) The most complete list is given by Sudlow, who describes the forms found in the Oudalan province of North-East Burkina Faso (the so-called *ša*- dialects, while these forms seem absent in the *za*- dialects) along with those attested in Mali (Sudlow 2009: 103, 111, 179, 233, 241):

(14)	masculine <i>əssənān</i> (ša), <i>issənān</i> (Mali)	feminine <i>tāssən-en</i> (ša), <i>tissən-en</i> (Mali)	“both”
	<i>ākkārāḍān</i> (ša), <i>ikkārāḍān</i> (M.)	<i>tākkāreḍān</i> (ša), <i>tikkārāḍen</i> (M.)	“all three”
	<i>ikkəḗān</i> (M.)		“all four” (The Malian forms reported here correspond to those recorded for “both” by Heath (2005: 256, 2006: 627): m. <i>i-ssən-æn</i> f. <i>t-i-ssən-en</i> . However, the latter does not record the forms for “the three” and “the four.” For the sake of uniformity, the slight difference in orthography (Heath’s <æ> vs. Sudlow’s <ā>) will be disregarded in the following examples: only <ā> will be consistently used.)

Another remarkable feature affecting definite numerals from “two” onward in Tuareg emerges when they are accompanied by a (plural) personal pronoun to express “all X of us/you/them.” In this case, they require a clitic pronoun belonging to a particular series, characterized by an initial vowel *e* or *i*, which is little used outside this construction (the preposition *gar* ‘between’ and a very limited set of nouns). Although the dialects of Niger seem devoid of special forms to express all of 2, 3, and 4,” they also share the use of a peculiar series of pronouns accompanying numerals, for example, *kəraḍ-ewən* “the three of you” (Air: Petites Sœurs 1974: 201). See also the forms for pronouns affixed to numerals 2–10 in Azawagh and Tayert: *eššin-enā(y)/-ewwān/-essān* ‘both of us/you/them’ (Prasse et al. 2003: 603, 727, 825).

Foucauld describes the phenomenon in his sketch of Tuareg grammar Foucauld (1920: 51–53): “the affixed pronoun that accompanies these words [i.e. cardinal numerals from 2 to 19, and the expressions meaning ‘all 2, 3, 4’] is always the same that accompanies *matt* ‘mothers’ (*īnāy*, *ewān*, *ekmāt*, *esān*, *esnāt*).” (The pronouns appearing in Foucauld’s passage are written out *īner*, *iouen*, *ikemet*, *isen*, and *isenet*. In the quotation, I have replaced them with the more accurate forms identified by Prasse in later works (Prasse 1993; Prasse et al. 2010).)

Given the peculiarities of these constructions, Heath (2005: 252) singles out a category of “pronominalized numerals” with reference to the special forms assumed

by the complex resulting through the insertion of the vowel *e* between the numeral and the pronominal affix.

- (15) “both of them” m. *əssin-e-ssān* f. *sānat-āt-e-snāt* (Heath 2005: 252)
 “the three of us” m. *ākkārāḍ-e-nāy* f. *ākkārāḍ-āt-e-nāy* (ibid.)

In the last example, it is noteworthy that the numeral preceding *-e-* in “the three of us” is not the “regular” form m. *kāraḍ* f. *kāraḍ-āt*, but a form with geminated *k*, like in the definite forms *ākkārāḍān* (ša), *ikkārāḍān* (M).

In Burkina Faso, the numeral preceding the pronominal affix displays an ending *-ān*, identical to that of the definite numerals described above, and this is not limited to “three” or “four” and also affects higher digits (Sudlow 2001: 339):

- (16) “two of them” *əssənān-essān*
 “three of us” *kāraḍān-enāy*
 “four of you” *əkkōḍān-ewwān*
 “seven of them” *əssayān-essān*

The examples in (16) come from the *za-* dialects of Oudalan. The corresponding feminine forms display the simple feminine numeral, devoid of the *-ān* appendix: “two of them (f.)” *sānatāt-esnāt*, etc. However, in the *ša-* dialects the definite numeral *əssənān* is used for both genders: “both of them” m. *əssənān-essān*, f. *əssənān-esnāt* (Sudlow 2009: 241).

Southern and Northern Morocco: *Tashelhit*, *Iznassen*

The Tashelhit dialects of Southern Morocco do not use *isnin* or similar expressions for defined numerals. But they share with Tuareg a particular way of linking numerals with plural personal pronouns. In fact, in such cases, a particle *-it-* is inserted between the numeral and the pronominal affix. This feature is ancient as it already appears in texts composed by al-Awzal (d. 1748). The following examples come from a grammar of the modern language (17.a) and a study on ancient texts (17.b).

- (17) a. *ḥtan s-sin-itsən* “the two of them left”
təffḃəm s-kraḍ-itun “the three of you went out” (Aspinion 1953: 259)
 b. *sin itsn* “the two of them (m.)” *snat itsnt* “the two of them (f.)”
id bu zzit kkuz-itsn “oleaginous grains, the four of them”
imawn n lāɛdab sa-itsn “the gates of Hell, the seven of them”
 (van den Boogert 1997: 287)

The northern varieties of the Rif region have lost almost all the Berber numerals and make use of Arabic loans from “two” onward. Consequently, they too lack a special numeral for “both.” However, in the eastern dialect of Iznassen, there exists a

word *snäin* “both” that looks similar to (*i*)*snin* described above (18.a). Renisio, the only researcher who records this form, writes it as one word within the texts but tries to explain it, in the grammatical part, as a prepositional phrase composed of the preposition *s* “from; with, etc.” + *näin*, abridged form of *ətnäin* “two” (1932: 213). When he wrote his grammar, *isnin* and cognates were not known, and the explanation he adopted tried to bring together this peculiar form with the existing knowledge about numerals in that language, while a link with *isnin* (admittedly, with a small adaptation to the form of *ətnäin*) cannot be ruled out.

Apart from this, the Iznassen variety shares with Tashelhit the peculiar way of placing an element *-it-* between the numeral and the pronoun in constructions meaning “the X of us/you/them” (18.b). This element appears as *-is-* in other varieties of the Rif region (18.c), (Despite the formal similarity, this element can hardly be identified with a form assumed by the preposition *s* (several meanings, among which “by means of,” “out of,” etc.) before a pronoun, which appears as *is* in some other Berber languages, because in the Rif region, it assumes other forms. Cp. *idjen zi-nay* “one of us” (Iznassen and Renisio 1932: 115).) as can be seen in the following examples:

- (18) a. *nəšin snäin anqarṛəš middən* “the two of us bite the people” (Renisio 1932: 169)
- b. *əmmuṭən snäin iṭsən* “they died both of them” (Renisio 1932: 167)
usin-d ɣlaṭa iṭsən “the three of them came” (Renisio 1932: 109)
lbæd n təḡbirin, səbæa yəṭsənt, bdant səssənt zi ɣiṭ “some doves – seven of them – began to drink from the water-hole” (Kossmann 2000: 130)
- c. *tfənn-anay sxəmsa is-nay* “they captured the five of us” (Renisio 1932: 267)
u bæḡdaha iḡyaq ɣ iṣḡiqən-ənnay slaṭa is-sən “later on, he released our three brothers” (ibid.)
nəqqim nəkkini d uṣḡiqinu s-əjjuj iss-nay “we remained, my brother and I, the two of us” (ibid.)

It cannot be excluded that a similar affix between numeral and pronoun originated the particle *id* that binds numeral and pronoun in other parts of the Berber world, for instance, in Kabyle, where it is usually considered as the comitative preposition “with”: *di sbæa id-sən* “the seven of them,” etc.

Zenaga

The Zenaga language of Mauritania is still little known as far as certain subtleties of grammar are concerned: the behaviour of “defined numerals” belongs here. A study focused on numerals in Zenaga, by Taine-Cheikh (2005), deals with several features of these parts of speech but does not take into consideration the way of expressing numerals when they are “defined” (“the X,” “all the X,” “the X of us,” etc.). She includes these expressions in her vast dictionary under the entries “two,” “three,”

and “four” (Taine-Cheikh 2008: 488, 310, 306). Previously, this subject has been briefly treated by Nicolas (1953), a text hard to approach for its lack of organicity and numerous inconsistencies, along with the cumbersome spelling adopted by the author, not to mention many typos. (According to Taine-Cheikh (2008: LVI ff.), the real author was, most probably, Mokhtar Ould Hamidoun, a native scholar. In the quotations, I keep the original spellings, with the only exception of the spirantization mark, which Nicolas writes out with a caron above the letter, and I replace with a short line below.) In a short paragraph devoted to numerals (Nicolas 1953: 62), this text records:

- | | | |
|------|------------------------------------|------------------------------|
| (19) | “both” (“tous les deux”) | <i>tənsnè'd šənh, təšnih</i> |
| | “all three” (“tous les trois”) | <i>kārd è'd šənh</i> |
| | “all four” (“tous les quatre”) | <i>akkuṭ è'd šənh</i> , etc. |
| | “the two of us” (“à nous deux”) | <i>a'nəgh təšnīn</i> |
| | “the three of us” (“à nous trois”) | <i>a'nəgh kārd è'd nəgh</i> |

The only occurrence of one of these numerals in context, within the sample of texts contained in the book, is *əfəðssə-əš təšnīh* “both his hands” (Nicolas 1953: 73).

In this list, clearly only *təšnīh* means “both” without accompanying pronouns, while all the other examples provided by Nicolas refer to syntagms meaning “all X of us/them” and consisting of numeral + *è'd* + pronouns. In these constructions, one can notice the presence of a particle *è'd* that recalls the element inserted between numerals and pronouns under the same conditions in Tashelhit (-it-) and Tuareg (-e/-i-). According to Taine-Cheikh (2008: 488, 310, 306), the phonetic shape of the segment, inserted in such constructions, is *-əḏd-*:

- | | | |
|------|-----------------------------------|---|
| (20) | “the two of them” (“eux deux”) | <i>nəhni təšn – əḏd-šən (< təšnīh əḏd-šən)</i> |
| | “the two of us” (“nous deux”) | <i>nəkni təšn – əḏd-nəg (< təšnīh əḏd-nəg)</i> |
| | “the two of you” (“vous deux”) | <i>nətni təšn – əḏd-kūn (< təšnīh əḏd-kūn)</i> |
| | “the three of them” (“eux trois”) | <i>nəhni kaṛd – əḏd-šən (< kaṛad əḏd-šən)</i> |
| | “the three of us” (“nous trois”) | <i>nəkni kaṛd – əḏd-nəg (< kaṛad əḏd-nəg)</i> |
| | “the three of you” (“vous trois”) | <i>nətni kaṛd – əḏd-kūn (< kaṛad əḏd-kūn)</i> |
| | “the four of them” (“eux quatre”) | <i>nəhni akkuṭ – əḏd-šən (< akkuṭ əḏd-šən)</i> |
| | “the four of us” (“nous quatre”) | <i>nəkni akkuṭ – əḏd-nəg (< akkuṭ əḏd-nəg)</i> |
| | “the four of you” (“vous quatre”) | <i>nətni akkuṭ – əḏd-kūn (< akkuṭ əḏd-kūn)</i> |

According to Taine-Cheikh, the alternative form of “two” appearing in Nicolas’ list, *təšnīh* (other spellings *təšnīh*, *təšnīn* and *tənsn-*), is a feminine noun used only in two circumstances: (i) preceded by the preposition *ən* “of” to express the ordinal “the second (m. and f.)” and (ii) in the above-mentioned constructions meaning “both.” The initial *tə-* is a marker of feminine nouns, and the word could be interpreted as an

abstract name like “the couple” (not simply “couple” because, anyway, in both instances it is semantically “defined”).

It is worth noting that in Zenaga the “regular” number “two” is also somehow problematic as Taine-Cheikh (2005: 271) pointed out:

“The example of ‘two’ is particularly instructive because this cardinal has two different forms depending on whether it is used alone (*šinān*) or in construction (*šin* [*nʷ*]). While the second form presents some difficulties of analysis (...), the comparison of *šinān* with the form (F) *šinād* of Zenaga or with those of other dialects (e.g., *sən* of Ghadamsi) encourages us to see, in *šinān*, the form reinforced (expressive) of the cardinal.” (“L’exemple de ‘deux’ est particulièrement instructif car ce cardinal présente deux formes différentes selon qu’il est employé seul (*šinān*) ou en construction (*šin*[*ny*]). Si la seconde forme présente quelques difficultés d’analyse (...), la comparaison de *šinān* avec la forme (F) *šinād* du zénaga ou avec celles des autres parlers (ex. *sən* du ghadamsi) incite à voir, dans *šinān*, la forme renforcée (expressive) du cardinal.”)

According to Rössler (1952: 142), the long form of the numeral is the proof of a triconsonantal root: “the Libyan form is evidently based on a dual of a noun from the root *š-n-n*: **šinānā*, f. **šinānatā*.” (“Der libyschen Bildung liegt offenbar ein Dual eines Nomen von der Wurzel *š-n-n* zugrunde: **šinānā*, f. **šinānatā*.”) This reconstruction was based on the only description existing at that time, namely, R. Basset’s *šinān* f. *šenānet* (1909: 218). According to Taine-Cheikh (2008: 488), the form *šinān* is “pronominal” (i.e., not accompanied by a noun) and used for both genders, but Nicolas (1953: 60) concords with R. Basset in recording two forms: *šənən^h* and f. *šənanəṭ*. It seems that the modern language has lost the feminine form. (As a matter of fact, Nicolas (1953: 119), himself, records two diminutive forms, with a masculine and a feminine prefix, both affixed to the masculine form *šənən^h*. Moreover, he records only the “masculine” form of “two” for use with nouns of both genders, as appears in *šən tšūggaⁿh* “two mats” (Nicolas 1953: 61), while Taine-Cheikh (2008: 488) records two forms: m. *šin/šin^y* and f. *šināt*.)

Another tentative explanation for the longer form of this numeral has been put forth by Blažek (1999: 62), who believes that “the final extension in *-an/-ən* in Zenaga corresponds to the collective of other Berber languages,” that is, the ending *-in* of *isnin* and cognates.

The *Kitāb al-Barbariyya*

The recent rediscovery of some manuscripts containing a huge medieval text on religious matters entitled *Kitāb al-Barbariyya* opened the way to a better knowledge of an archaic stage of Berber. (On this text, see Ould-Braham (2008) and Brugnatelli (2016); on its language, see Brugnatelli (2011, 2014).) One of the most noteworthy features of this language is the existence of a rich series of defined numerals. The expression *isnin* is well attested, as shown in example (21): (The following quotations refer to the page and line numbering of the manuscript MS.ARA.1936 of the Bulac Library in Paris.)

numeral that originally contained the velar sound *[k], the typical palatalization of the Zanata languages has two different reflexes:

- > š [ʃ] at the beginning of the word: **karəḍ* > *šarəḍ* “three.”
 > y [j] when an initial vowel precedes it: **əkard-in* > *əyard-in* “the three.” (About the reflexes š and y < *k in the Zanata languages, see Kossmann (1999: 173 ff.). A general rule for the change *k > y is difficult to find (ibid. 186 ff.), although it never happens at the beginning of a word, and in many cases y develops after the prefixation of the nominal marker *a-* or *ta-/ti-* (“Noms avec y dérivés de verbes avec k”, ibid. 192–193).)

Diachronic Considerations

The survey made in the first part of this paper aimed at pointing out several expressions used throughout the Berber-speaking territory to express “defined numerals.” Most of them are limited to the numeral “two,” but in some instances longer series could be detected.

The following table summarizes the forms used for “both” in some Berber languages, along with the special forms for “the three” in the languages that possess them:

		masculine	feminine
B. Menacer		<i>snin</i>	
Aurès		<i>isnin</i>	<i>tisnin/isntin</i>
	“the 3”	<i>itlatəggən</i>	<i>titlatəggin</i>
Kabyle		<i>isnin</i>	<i>tisnin (?)</i>
	“the 3”	<i>i ɣlaɣa</i>	=
Zuara		<i>isnin</i>	<i>tisəntin</i>
El Qalaa		<i>isnin</i>	<i>isnintən</i>
Jerba		<i>isnin</i>	<i>tisəntin</i>
Mzab		<i>i/ay snin</i>	<i>ay sənnətin</i>
	“the 3”	<i>ay šarɖin</i>	<i>ay šarəɖtin</i>
Tuareg (North)		<i>əssənān</i>	<i>(t)əssənāten</i>
	“the 3”	<i>əkkərɖān</i>	<i>(t)əkkərɖāten</i>
Tuareg (South)		<i>əissənān</i>	<i>tā/tissənen</i>
	“the 3”	<i>ā/ikkārāɖān</i>	<i>tākkārəɖān/tikkārāɖen</i>
Zenaga		<i>təšnih</i>	=
K. Barbariyya		<i>isnin</i>	<i>tisəntin</i>
	“the 3”	<i>əyyarɖin</i>	<i>təyyarɖin</i>

The survey also includes the special ways of joining numeral and pronoun in the expressions meaning “(all) the X of us/you/them,” in which a “weak” type of definiteness is involved. In such cases, many Berber languages display an element inserted between the numeral and the pronoun: Tuareg *-e-*, Tashelhit *-it-*, Rif *-it-/is-*

Zenaga *äʔd-*, and possibly also Kabyle *id/yid*. This feature closely reminds of the behaviour of plural pronouns suffixed to kinship terms in most Berber languages, because, here too, an element of uncertain origin (*-t-*) is inserted between noun and affix (for instance, *baba-s* “his father” but *baba-t-sen* “their father”). Both constructions share a trait of (“weak”) definiteness and one is tempted to see in them a way of expressing this feature. It would be interesting to inquire into the diachronic origin of the morphemes that characterize such constructions, but it would be a long and highly speculative research, exceeding the scope of the present paper.

Coming back to the “strongly” defined numerals summarized in the above table, their grammatical status and diachronic origin are disputed. Looking for an explanation of *m. isnin f. tisin* in the Aurès, A. Basset puts forward the hypothesis that they are the relics of ancient verbal forms, displaying an invariable participial ending *-in* and preceded by a “demonstrative pronoun” masculine plural *i*, feminine *ti*. (Basset (1945–1948: 19): “on est tenté de retrouver à la finale une désinence de pluriel participiale (*-in* commun), et à l’initiale un pronom démonstratif m.p. *i*, f.p. *ti*, suivant une formule usuelle en touareg.”) The verbal nature of these forms is assumed based on the analysis he made of the Tuareg defined numerals (ironically classified among the “undefined pronouns” by Foucauld (1920: 103–104)): “they seem ancient 3rd pers. m.pl. of verbs that would be to *ssin* ‘two’ and *kraq* ‘three’ as *mru* ‘be ten’ is to *mraw* ‘ten’” (Basset 1945–1948: 19). As a matter of fact, a typical feature of Tuareg is the extensive use of verbs in cases where other Berber languages would rather use nominals: for instance, instead of using adjectives, Tuareg expresses qualities through “quality verbs.” (The hypothesis of “a probable verbal origin” is shared by Chaker (2012: 5612), who limits himself to resume Basset’s conclusions.) In support of this analysis, Basset invokes the existence of a verb *sakkəraq* “to triple, do three times, do the third time.” (Prasse et al. (2003): “tripler, faire trois fois faire pour la troisième fois.” The meaning of this verb in Tahaggart, quoted by Basset, is “poser trois pions dans trois cases contiguës.” Basset also evokes a Kabyle form *sənnunūt* “repeat (a word),” which appears only once in an ancient poem and has never been found elsewhere in Kabyle. Apropos he writes: “it provides a verbal form in relation -we think- to the numeral concerning ‘two.’”) The form *sakkəraq* is characterized by the prefix *s-* of “causatives” shared by many Berber languages to derive verbs from nouns (cp. Jerba *skufəs* “to spit” < *ikuḥsan* “spit,” *skərkəs* “to lie” < *tikerkas* “lie”). The Tuareg form is *səBBəCəD*, which “corresponds to prefix-less *nomina actionis* listed under IV.K.4.b. Mostly, the simple verbs from which these *nomina actionis* derive are not in use” (Prasse 1973: 58) (This group of *nomina actionis* is described in Prasse (1974: 87–88). Forms like *əssənə* and *əkkərdən* might belong here, like *əbbələl*, pl. *əbbələlān*, infinitive of *bālāl* “have everything in abundance,” or *əmməḡəd*, pl. *əmməḡədān*, infinitive of *māḡād* “be a plebeian, a vassal.”).

Even if a remote verbal origin of numerals cannot be excluded, from a synchronic point of view the verbal nature of the defined numerals in Tuareg is far from established. In his description of Tahaggart, Prasse calls them “collective numerals” and remarks: “They are visibly formed on the same root as the corresponding cardinal numerals, but with a different vocalization (**ī-ī*? Note the group *rɔ* in

akkərḍən, resulting from the dropping the characteristic vowel?). They apparently display an ending *-ən* ([. . .] f. *-în* as in plural #1), that of feminine however, unexpectedly, preceded by *-āt* (not *-t*) of participles and of the qualitative perfect (as well as of isolated personal pronouns pl. f. *näkkānātiḍ* etc.). The ending seems to me to correspond to *-in* (**-īn*) of Northern Berber and is perhaps an abbreviated form of it.” (“Ils sont visiblement formés sur la même racine que les noms de nombre cardinaux correspondants, mais avec une vocalisation différente (**ī-ī?* noter le groupe *rḍ* dans *akkərḍən*, résultat de la chute de la voyelle caractéristique?). Ils se terminent apparemment sur une désinence *-ən* (ainsi Gb.Al.: f. *-în* comme au pl.1), celle du f. étant cependant, chose inattendue, précédée de *-āt* (non *-t*) des participes et des pf. qualitatifs (et aussi des pronoms pers. isolés pl. f. *näkkānātiḍ* etc.). La désinence me semble correspondre à *-in* (**-īn*) du BN et en est peut-être une forme abrégée” (Prasse 1974: 408).)

Whatever its origin, the puzzling *-āt-* of feminine is also the same ending of the corresponding standard numerals, and this might suggest that the defined numerals are built upon them, rather than starting from a verbal root.

As far as Malian Tuareg is concerned, Heath (2005: 256) compares the endings of the definite numerals to the suffixes of participle but remarks that “the relationship between the masculine and feminine is irregular, especially since [the former] ends in the Masculine Singular Participial suffix *-ān* while [the latter] ends in Plural Participial suffix *-en*.” However, turning to the morphology of participles for a comparison seems hardly appropriate. Letting aside the fact that the participial ending in the plural is *-nen* and not *-en*, a straightforward parallel is offered by the nominal plural suffixes. In fact, the plural endings of nouns are precisely *-ān* in the masculine and *-en* in the feminine. If one considers that nouns also have prefixes identical to the ones of these numerals (i.e., masculine *i-* and feminine *ti-*), the defined numerals appear, morphologically built, as plural nouns: masculine *i-ssən-ān*, *i-kkārāḍ-ān*, feminine *ti-ssən-en*, *ti-kkārāḍ-en* “the two, the three,” just like *i-zew-ān* “twigs” (m.) versus *ti-zew-en* “small twigs” (f.).

The medieval forms offer support to the hypothesis of nominal morphemes attached to standard numerals. In particular, the feminine form *tisəntin* (still retained as such in Zuarā and Jerba) clearly appears built on the cardinal number *sənt*, with the prefixation of the marker of plural feminine nouns *ti-*, and the suffixation of *-in*. The former morpheme comes from an old demonstrative article, (There is a rich literature on the subject of Berber demonstratives that became nominal markers passing through a stage of prefixed definite article. See Vycichl (1957) and Brugnatelli (1997, 2006).) while the latter is a postposed demonstrative that can be found in some personal pronouns (like *šəkkīn* “you, thou,” emphatic form of *šəkk*) or in expressions like *w-in* “that” (vs. *w-u* “this”) and *əwrin* “elsewhere” (cp. Moroccan Tamazight *iwər-in(n)* “over there”). Above all, this deictic appears in the particle *-din* that is postposed to nouns in anaphora. The parallel between the use of *-din* in nouns and *-in* in numerals is best exemplified in (23.a). Nondefined numerals, on the contrary, retain remnants of an archaic construction with nouns devoid of the initial syllable and of any plural marker (23.b) as an alternative to the “regular” one:

- (23) a. *tiferni n wuḍmawen-din eyarḍ-in* [Arabic gloss: *al-talāṭa*]
 “the choice of (among) the three ways (of acting against a murderer:
 death, pardon or payment of blood-money)” (f. 201b, l. 3)
- b. *sen yil* “two cubits”
cared lyem “three camels”
uqqez yur “four months”

Conclusions

To sum up, it seems that the defined numerals, still in use in some Berber languages, trace back to an archaic stage when the ancient demonstratives that eventually coalesced with nouns and lost any trait of definiteness still retained a value of definite articles. Over time, these morphemes have simply become a nominal mark in most nouns, while in numerals they gave rise to a series of defined numerals. These have undergone some minor modifications but still appear in some modern languages: up to “four” in Tuareg and up to “ten” in Tashawit and in Tumzabt: while *isnin* “both” is also in use in several other languages, in particular in Jerba, Zuara, and Tripolitania.

Unlike standard numerals, defined numerals mostly display typically nominal morphemes. They are always treated as nominals: either alone or after a name or a pronoun as an apposition. The nominal character of defined numerals is also stressed by the possibility of being further determined through the use of demonstratives, as is the case in some of the examples already quoted: *yih isnin* “these two” (8), *ay sennin-u* “these two” (12), *ayyardin-uḍa* “these three (witnesses)” (22). In this respect, defined numerals, which are usually limited to lower numbers, do not fit into the general principle suggested by Corbett (1978: 355): “nouniness increases with numerical value.” This is not surprising, given that, owing to their peculiar morphology, they can no longer be considered “simple cardinal numerals” in the strict sense.

From a broader typological point of view, it is worth noting that in many languages, the category of “defined numerals” can be expressed not only through standard strategies of definiteness but also through dedicated lexical items, although this feature is mostly limited to the expression of “both” (For example Italian *entrambi* and *ambedue*, German *beide*, Russian *оба* (*oba*), Greek *ἄμφω* (*ámphō*), Sanskrit *उभ* (*ubha*), Arabic *كِلَا* (*kilā*), Japanese *両方* (*ryōhō*)) and is much rarer in the case of higher numbers. This is probably in connection with the frequent occurrence of “natural” duals, that is, item pairs (body parts, objects made up of two symmetric parts) that call for a synthetic reference, while groups of three or more items are mostly “occasional,” (Terminology of Fontinoy (1969: 14 ff.)) and their occurrence is less frequent. The fact of possessing peculiar ways of expressing defined numerals higher than “both” makes Berber languages an interesting field of observation for further research on the phenomena related to this grammatical category.

References

- Aspinion R (1953) Apprenons Le Berbère: Initiation aux dialectes chleuhs. F. Moncho, Rabat
- At-Abbas (1976) Contes merveilleux et fables. Textes nouveaux dans le parler des At-Abbas. le Fichier Périodique, Alger
- Aulard (1989) La Vie dans le Mزاب. Textes bilingues (1938–1941). Études Et Documents Berbères 5:131–168. <https://doi.org/10.3917/edb.005.0131>
- Basset A (1945) Berbère isnin ‘tous les deux’. C R Groupe Linguist Etud Chamito-Sémitiques 4: 19–20
- Basset A (1961) Textes berbères de l’Aurès (Parler des Aït Frah). Aderien-Maisonneuve, Paris
- Basset A (1969) La Langue berbère. Published for the International African Institute by Dawson of Pall Mall, London [reprint of the 1952 edition]
- Basset R (1884) Notes de lexicographie berbère. Le dialecte des Beni Menacer. J Asiat 518–556: 148–219
- Basset R (1890) Loqmân berbère. Leroux, Paris
- Basset R (1892a) Notice sur les dialectes berbères des Harakta et du Djerid tunisien. Working: Oriental University institute [publications du Neuvième Congrès International des Orientalistes – Londres 1891, 18 p]
- Basset R (1892b) Textes berbères dans le dialecte des Beni Menacer. G Soc Asiat Ital 6:37–84
- Basset R (1909) Mission au Sénégal. Tome I: Étude sur le dialecte zénaga – Notes sur le hassania – Recherches historiques sur les Maures. Leroux, Paris
- Beguino F (1917–1918) Il gergo dei berberi della Tripolitania. Ann Ist Orient Napoli:107–112
- Bekkay S (2012) Tanfust n sat tziwin d Muggayla – L’histoire des sept filles et l’ogresse. In: Nouh-Mefnune B (ed) Tinfas: contes en berbère mozabite. Çra n yiḍrisen d tinfas s Tugyasant. eBox éditions, pp 20–24. <http://www.eboxeditions.com>
- Berkai A (2012–2013) Essai d’élaboration d’un dictionnaire Tasaḥlīt (parler d’Aokas)-français. Thèse de doctorat, Université Mouloud Mammeri de Tizi-Ouzou
- Blažek V (1999) Numerals: comparative-etymological analyses of numeral systems and their implications (Saharan, Nubian, Egyptian, Berber, Kartvelian, Uralic, Altaic and Indo-European languages). Masarykova Univerzita, Brno
- Bossoutrot A (1900) Vocabulaire berbère ancien (dialecte du djebel Nefoussa). Revue Tunisienne 7: 489–507
- Brugnatelli V (1982) Questioni di morfologia e sintassi dei numerali cardinali semitici. La Nuova Italia, Firenze
- Brugnatelli V (1997) L’état d’annexion en diachronie. In: Bausi A, Tosco M (eds) Afroasiatica Neapolitana. Contributi presentati all’8° Incontro di Linguistica Afroasiatica (Camito-Semitica) – Napoli 25–26 Gennaio 1996. I. U. O, Napoli, pp 139–150
- Brugnatelli V (2006) L’ancien “article” et quelques phénomènes phonétiques en berbère. In: Ibrizimow D, Vossen R, Stroemer H (eds) Etudes berbères III. Le nom, le pronom et autres articles. Actes du “3. Bayreuth-Frankfurter Kolloquium zur Berberologie”, 1–3 juillet 2004. Köppe, Köln, pp 55–70
- Brugnatelli V (2011) Some grammatical features of ancient eastern Berber (the language of the *Mudawwana*). In: Busetto L (ed) He bitaney lagge. Dedicato a/dedicated to Marcello Lamberti. Saggi di Linguistica e Africanistica – essays in linguistics and African studies. Qu.A.S.A.R., Milano, pp 35–46
- Brugnatelli V (2014) Typology of eastern medieval Berber. STUF Lang Typol Univ 67(1):127–142
- Brugnatelli V (2016) Un témoin manuscrit de la “*Mudawwana* d’Abū Ḡānim” en berbère. Étud Doc Berbères 35(36):149–174
- Chaker S (2012) Nombres/numération. Encycl Berbère 34:5610–5616
- Cheng LL-S, Heycock C, Zamparelli R (2017) Two levels for definiteness. In: Erlewine MY (ed) Proceedings of GLOW in Asia XI, vol 1. MIT Press, Cambridge, MA, pp 79–93. [MIT working papers in linguistics, vol 84]. <https://glowlinguistics.org/asia11/wp-content/uploads/sites/3/2017/11/v1-cheng-heycock-zamparelli.pdf>

- Corbett GG (1978) Universals in the syntax of cardinal numerals. *Lingua* 46:355–368
- Dallet J-M (1982) Dictionnaire kabyle-français. In: *Parler des At Mangellat (Algérie)*. SELAF, Paris
- de Foucauld C (1920) Notes pour servir à un essai de grammaire touarègue (dialecte de l'Ahaggar). Carbonel, Alger
- de Foucauld C, de Calassanti-Motyliniski A (1922) Textes touareg en prose (Dialecte de l'Ahaggar): publiés par René Basset. Carbonel, Alger
- Delheure J (1984) *Ağraw n yiwalen tumzabt-t-tfransist*. Dictionnaire mozabite-français. SELAF, Paris
- Fontinoy C (1969) Le duel dans les langues sémitiques. Les Belles Lettres, Paris
- Galand L (1967) La construction du nom de nombre dans les parlers berbères. In: *Verhandlungen des 2. Internationalen Dialektologenkongresses*. Marburg/Lahn 5–10. Sept. 1965. Steiner, Wiesbaden, pp 253–259. [ZMaF, Beihefte 3–4]
- Galand L (1969) Types d'expansion nominale en berbère. *Cahiers Ferdinand Saussure* 25:83–100
- Gutova E (2011) The Sanusi Creed in Kabyle Berber: manuscript Ka 21 from the Lmuhub Ulahbib Library (Béjaïa, Algeria). Master dissertation, Leiden University
- Gutova E (2016) The Sanusi Creed in Kabyle Berber: manuscript Ka 21 from the Lmuhub Ulahbib Library (Béjaïa, Algeria). *Étud Doc Berbères* 35(36):263–294
- Heath J (2005) A grammar of Tamashek (Tuareg of Mali). Mouton, Berlin
- Heath J (2006) Dictionnaire touareg du Mali: tamachek-anglais-français. Karthala, Paris
- Iwalen – dicco* (n.d.) Mozabite-French dictionary in pdf, devoid of title, author and date. <http://tumzabt.free.fr>. 571 p. Last visited 10 Jan 2021
- Kossmann MG (1999) Essai sur la phonologie du proto-berbère. Köppe, Köln
- Kossmann MG (2000) A study of eastern Moroccan Fairy Tales. Academia Scientiarum Fennica, Helsinki
- Lafkioui M, Merolla D (eds) (2002) Contes berbères chaouis de l'Aurès, d'après Gustave Mercier. Köppe, Köln
- Mercier G (1900) Mœurs et traditions de l'Aurès. Cinq textes berbères en dialecte chaouia. *J Asiat* 9 (xvi):189–248
- Mitchell TF (2009) Zuaran Berber (Libya). Grammar and texts. Köppe, Köln
- Mouliéras A (1999) Contes merveilleux de Kabylie. Edisud, Aix-en-Provence
- Nicolas F (1953) La langue berbère de Mauritanie. IFAN, Dakar
- Ould-Braham O (2008) Sur un nouveau manuscrit ibāḍite-berbère. *La Mudawwana d'Abū Ḡānim al-Ḥurāsānī* traduite en berbère au Moyen Âge. *Étud Doc Berbères* 27:47–71
- Penchoen TG (1973) Étude syntaxique d'un parler berbère (Ait Frah de l'Aurès). I.U.O., Napoli. [Studi Magrebini 5]
- Petites Sœurs de Jésus (1974) Contes touaregs de l'Aïr. SELAF, Paris
- Prasse K-G (1973) Manuel de grammaire touarègue (tāhāggart). VI–VII Verbe. Université de Copenhague, Copenhague
- Prasse K-G (1974) Manuel de grammaire touarègue (tāhāggart). IV–V Nom. Université de Copenhague, Copenhague
- Prasse K-G (1984) Compte rendu de Dallet (1982). *Acta Orient* 45:168–172
- Prasse K-G (1993) Du nouveau sur la vocalisation de la *tahaggart*. In: Drouin J, Roth A (eds) *À la croisée des études libyco-berbères*. Mélanges offerts à Paulette Galand-Pernet et Lionel Galand. Geuthner, Paris, pp 269–285
- Prasse K-G, Alojali G, Mohamed G (2003) Dictionnaire touareg-français (Niger), 2 vols. Museum Tusculanum Press, Copenhague
- Prasse K-G, de Foucauld C, de Calassanti-Motyliniski A (eds) (2010) Textes touaregs en prose (Dialecte de l'Ahaggar) Publiés par René Basset Alger 1922. Retranscription complète par K.-G. P. Det Kongelige Danske Videnskabernes Selskab, Copenhague
- Renisio A (1932) Étude sur les dialectes berbères des Beni Iznassen, du Rif et des Senhaja de Sraïr. Grammaire, textes et lexique. Leroux, Paris
- Rößler O (1952) Der semitische Charakter der libyschen Sprache. *Z Assyriol* 50:121–150
- Stumme H (1900) Märchen der Berbern von Tamazratt in Südtunesien. Hinrich, Leipzig

- Sudlow D (2001) *The Tamasheq of North-East Burkina Faso*. Köppe, Köln
- Sudlow D (2009) *Dictionary of the Tamasheq of North-East Burkina Faso*. Köppe, Köln
- Taine-Cheikh C (2005) Les numéraux en zénaga. Contribution à la syntaxe des noms de nombre en berbère. *Studi Magrebini* 3:269–280. [Studi Berberi e Mediterranei. Miscellanea offerta in onore di Luigi Serra (2006)]
- Taine-Cheikh C (2008) *Dictionnaire zénaga-français*. Köppe, Köln
- van den Boogert N (1997) *The Berber literary tradition of the sous*. Nederlands Instituut voor het nabije Oosten, Leiden
- Vycichl W (1957) L'article défini du berbère. In: *Mémorial André Basset*. Adrien-Maisonneuve, Paris, pp 139–146

Section IV

Morphology and Syntax (and Typology)

Contents

Introduction	362
The Morphophonological Properties of Q-verbs in Taqbaylit Berber	362
Vocalic Properties	363
Consonant Gemination in the Perfective of Basic Q-verbs	367
Adjective Formation	367
The Morphosyntactic Properties of the Perfective Stem of Q-verbs	368
Agreement in the Perfective	369
The Interpretation of the Perfective	370
The Effect of Directional –d	371
The Formal Properties of Verbal and Nominal Templates	372
The ‘Normal’ Verbal Template	373
The Nominal Template	375
The Template of Q-verbs	379
Conclusion and Outlook	382
References	383

We would like to thank Amazigh Bedar for discussions and detailed comments on an earlier version of this article. We are grateful to the editors, Karim Bensoukas and Alireza Korangy, for inviting us to contribute to this volume, and to the reviewers for their comments.

S. Bendjaballah (✉)

LLING-UMR6310, CNRS & Nantes Université, Nantes, France

e-mail: sabrina.bendjaballah@univ-nantes.fr

M. Haiden

LLING-UMR6310, Nantes Université & CNRS, Nantes, France

e-mail: martin.haiden@univ-nantes.fr

Introduction

Berber verbs of state (“*verbes de qualité*”, henceforth “Q-verbs”) are well known for their peculiarities, which are, for the most part, well delimited and solidly described. Yet, the status of these particularities, in the grammar of the languages that display them and their implications for the theory of grammar, are still far from consensual.

In particular, the perfective paradigm of Q-verbs stands out in two respects: it is characterized by a reduced agreement paradigm and it has the interpretation of a pure state (e.g., *ɣif* “he is light”). In all other paradigms, Q-verbs have a normal, full inflectional paradigm, and they express a process, or a change of state. In this sense, the perfective of Q-verbs has long been considered to be related diachronically to the category of nouns. In this article, we explore the possibility of a nominal analysis of Q-verbs in synchronic grammar.

The chapter is organized as follows: in the first section, we introduce the morphophonological peculiarities of Taqbaylit Q-verbs. Then, we turn to their morphosyntactic properties. We examine their templatic structure, and compare it with that of verbs and nouns. The final section concludes the article.

The Morphophonological Properties of Q-verbs in Taqbaylit Berber

A representative set of the different Taqbaylit verb types is given in (1a). Each verb has 4 stems, which are used in different temporal/aspectual contexts. (1b) illustrates the derivation of the causative by prefixation of *s-*. In (1a), the intensive is formed in different ways: gemination and prefixation. (The prefix is realized as short [t^s] before a CV-initial stem, and as a geminate before V-initial stems and geminate initial stems.) In (1b), intensive formation resorts to stem-internal vowel alternations only: the causative prefix and the intensive prefix are in complementary distribution.

(1) a. Base

	<i>base (aorist)</i>	<i>intensive</i>	<i>perfective</i>	<i>negative perfective</i>	
i.	xðəm	xəddəm	xðəm	xðim	<i>work</i>
ii.	ɣli	ɣəlli	ɣli	ɣli	<i>fall</i>
iii.	ffəɣ	tt ^s əffəɣ	ffəɣ	ffɪɣ	<i>go out</i>
iv.	afəj	tt ^s afəj	ufəj	ufɪj	<i>fly</i>
v.	raði	t ^s raði	ruða	ruða	<i>agree</i>
vi.	faθ	t ^s faθ	fuθ	fuθ	<i>miss</i>

b. Derived causatives

	<i>base (aorist)</i>	<i>intensive</i>	<i>perfective (=neg. perfective)</i>	
i.	ssəxðəm	ssəxðam	ssəxðəm	<i>make work</i>
ii.	ssəɣli	ssəɣlaj	ssəɣli	<i>make fall</i>

iii. ssufəy	ssufuy	ssufəy	<i>make go out</i>
iv. ssafəj	ssafaj	ssufəj	<i>make fly</i>
v. sraðɪ	sraðaj	sruða	<i>convince</i>
vi. sfaθ	sfaθaj	sfuθ	<i>make miss</i>

The major types of Taqbaylit Q-verbs are given in (2) along with their derived causatives. In (2a), intensive formation always resorts to prefixation with geminated [t^s] as expected before a vowel-initial stem. By contrast, as in ‘regular’ verbs, the derived causative always expresses the intensive via stem-vowel alternation (2b).

(2) a. Base

	<i>base (aorist)</i>	<i>intensive</i>	<i>perfective (=neg. perfective)</i>	
i.	ixfif	tt ^s ixfif	xfif	<i>be light</i>
ii.	iβriç	tt ^s iβriç	βərriç	<i>be black</i>
iii.	ihliw	tt ^s ihliw	hlaw	<i>be sweet</i>
iv.	izwiɣ	tt ^s izwiɣ	zəgg ^w ay	<i>be red</i>
v.	alqaj	tt ^s alqaj	lqaj	<i>be deep</i>
vi.	imlul	tt ^s imlul	məllul	<i>be white</i>

b. Derived causatives^a

i.	ssixfəf	ssixfif	ssaxfəf	<i>make light</i>
ii.	ssiβrəç	ssiβriç	ssaβrəç	<i>make black</i>
iii.	ssihləw	ssihliw	ssahləw	<i>make sweet</i>
iv.	ssizwəɣ	ssizwiɣ	ssazwəɣ	<i>make red</i>
v.	ssalqi	ssalqaj	ssalqi	<i>make deep</i>
vi.	ssimləl	ssimlul	ssamləl	<i>make white</i>

^aThe examples in iii. and v. include a glide-final root, w and j respectively. In the absence of a preceding nucleus, final /j/ is realized as the corresponding high vowel [i], while final /w/ surfaces as such. This results in an asymmetry in the realization of e.g. /ssihlw/ → [ssihləw] ‘make sweet’ and /ssalqj/ → [ssalqi], *[ssalqəj] ‘make deep’

The morphophonological behaviour of Q-verbs differs from that of normal verbs in three respects, which we now examine.

Vocalic Properties

Q-verbs are characterized by the presence of two vowels, which are consistently located in initial position (before the first root consonant) and in penultimate position (before the last root consonant) (see Bendjaballah 2006, for data and a review of these classes).

(3)						example	#
a.	i	C ₁	C ₂	i	C _{2/3}	iβriç <i>be/become black</i>	27
b.	i	C ₁	C ₂	u	C _{2/3}	ifsus <i>be/become light</i>	5
c.	a	C ₁	C ₂	a	J	alqaj <i>be/become deep</i>	4
d.	{i/a}	C ₁		{i/a}	C ₂	ižið <i>be/become sweet/smooth</i>	1
						ažaj <i>be/become heavy</i>	1

The two vowels are identical, except for a group of five Q-verbs that have two different vowels, with *u* in penultimate position (3b). The presence of *u* in these verbs can, thus, be ascribed to the consonantal context since labials are well-known for their affinity with *u* crosslinguistically. In addition, note that there is some variation, with speakers accepting the forms with two identical vowels *u-u*, and that in other Berber languages, these verbs have an *u-u* melody.) This suggests that Q-verbs are equipped with a single lexical vowel that identifies several positions in the template. Note that the verbs in (3b) have a special feature: they all include a labial consonant (*imlul* ‘be white’, *ismum* ‘be sour’, *ifsus* ‘be light’, *imsus* ‘be tasteless’, *imyr* ‘be big’).

The presence of two identical vowels in the stem is specific to Q-verbs: normal verbs never have this shape (1). (As argued by Bader (1984, 1985) in his seminal work, schwa in Taqbaylit Berber has epenthetic status. Note that the distribution of schwa in Taqbaylit however does not seem to be always predictable by the phonotactics: morphological structure does play a role, too. We leave this question open for further research. In the following, the term ‘vowel’ refers to the three non-epenthetic vowels of the language: *i*, *a* and *u*.) The behaviour of the vowels in Q-verbs suggests that they are not part of the root, but that they are morphological markers. We briefly review three phenomena in support of this assumption: a) the distribution of vowels in Q-stems, b) their behaviour with respect to apophony in the perfective stem, and c) the shape of Q-roots that allow an optional regular derivation.

a) The distribution of vowels in Q-stems. In a normal (non-Q) Berber verb, the stem-vowels remain stable across the paradigm: they may be changed by an apophonic derivation, but they do not disappear. Consider the example in (1a.v) and (1b.v): if a verb has two stem-vowels in the aorist, it will have two stem-vowels throughout its paradigm. The vowels of Q-verbs lack such stability.

Consider first the initial vowel. (4a) reveals that it is absent in the perfective stem of the base verb. It is present in all other forms: the aorist and the intensive of the base verb and all stems of the derived causative, including the perfective (4b). The perfective of the base Q-verb, thus, stands out in that it is the only stem without an initial vowel.

Consider now the penultimate vowel: it is realized as a full vowel in all forms of the base verb. By contrast, it is absent in the aorist and the perfective of derived causatives. In other words, in the aorist and in the perfective, the penultimate vowel is in complementary distribution with the causative prefix. This fact must be considered seriously because normal verbs never exhibit such a vowel/zero

alternation between the base and the causative (see (1b)). Finally, the penultimate vowel is always (= in the base and in the causative) present in the intensive stem.

(4) a.

<i>vowel</i>	<i>base.aor.</i>	<i>base.pf.</i>	<i>base.int.</i>
initial	+	-	+
penultimate	+	+	+

<i>base.aor.</i>	<i>base.pf.</i>	<i>base.int.</i>	
ixfif	xfif	tt ^s ixfif	<i>be light</i>
iβriç	βərriç	tt ^s iβriç	<i>be black</i>
ihliw	hlaw	tt ^s ihliw	<i>be sweet</i>
izwiɁ	zəgg ^w əɁ	tt ^s izwiɁ	<i>be red</i>
alqaj	lqaj	tt ^s alqaj	<i>be deep</i>
imlul	məllul	tt ^s imlul	<i>be white</i>

b.

<i>vowel</i>	<i>base.aor.</i>	<i>caus. aor.</i>	<i>caus. pf.</i>	<i>caus. int.</i>
initial	+	+		+
penultimate	+	-		+

<i>base.aor.</i>	<i>caus. aor.</i>	<i>caus. pf.</i>	<i>caus. int.</i>	
ixfif	ssixfəf	ssaxfəf	ssixfif	<i>be light</i>
iβriç	ssiβrəç	ssaβrəç	ssiβriç	<i>be black</i>
ihliw	ssihləw	ssahləw	ssihliw	<i>be sweet</i>
izwiɁ	ssizwəɁ	ssazwəɁ	ssizwiɁ	<i>be red</i>
alqaj	ssalqi ^a	ssalqi	ssalqaj	<i>be deep</i>
imlul	ssimləl	ssamləl	ssimlul	<i>be white</i>

^aFinal *i* in *ssalqi* is the surface realization of root-final *j*. As such, it is expected to remain stable (cf. Bendjaballah 2001)

b) Apophony. In normal verbs, apophony systematically applies, where it can, between the aorist and the perfective, both in the basic form and in the derived causative, as can be seen in (1). By contrast, only some Q-verbs show the expected *i* → *a* alternation between the aorist and the perfective.

(5) and (6) sum up the apophony patterns observed in Q-verbs. Two restrictions can be observed. First, only *i* → *a* is observed, never *a* → *u*. This holds true both of basic verbs, and of derived causatives. This observation can be related to a distributional generalization concerning the Q-verbs with *a*-melody: these verbs all have $R_3 = j$ (*alqaj* ‘be deep’, *alwaj* ‘be slow/loose’, *aylaj* ‘be expensive’, *asxaj* ‘be generous’, *aɣaj* ‘be heavy’). In addition, no verb with $R_3 = j$ has a *i*-melody. There is thus a full correlation between the *a*-melody and the presence of *j* in R_3 . We propose that *j* in R_3 prevents the realization of a *i*-vocalization, and A, the apophonic output of I, surfaces instead. Apophony is never applied twice in the Berber verb system. Therefore, A cannot be apophonized to U, and *a* must not alternate. The peculiar behaviour of *a* in Q-verbs can thus be accounted for on phonological grounds.

The second restriction correlates with a morphological context. In the basic Q-verbs (5), *i* does not systematically apophonize between the aorist and the perfective: the expected *i* → *a* alternation is observed in 11 verbs and is absent in 17 verbs. By contrast, the expected *i* → *a* alternation is systematically obtained in the derived causatives (6). Note that in basic forms, the vowel is located in the penultimate position, while in the causative it is located in the initial position.

(5) Apophonic behaviour of basic Q-verbs^a

	vowel	aor.	pf.		
a. — apophony	<i>a</i>	alqaj	lqaj	<i>be deep</i>	4
		azaj	zaj	<i>be heavy</i>	1
	<i>i</i>	ixfif	xfif	<i>be light</i>	10
		iβriç	βərriç	<i>be black</i>	6
		izið	zizð	<i>be sweet</i>	1
b. + apophony	<i>i</i>	ihliw	hlaw	<i>be sweet</i>	4
		izwiʁ	zəgg ^w aʁ	<i>be red</i>	7

^aIn the *imlul-məllul* group, *u* does not alternate. This can correspond either to a non-apophonic pattern or to an instantiation of the *u* → *u* step in the Apophonic Path in the sense of Guerssel and Lowenstamm (1996). Because of this ambiguity, this group is absent in (5)

(6) Apophonic behaviour of derived causatives

	vowel	aor.	pf.		
a. — apophony	<i>a</i>	ssalqi	ssalqi	<i>make deep</i>	4
		ssaʒi	ssaʒi	<i>make heavy</i>	1
b. + apophony	<i>i</i>	ssixfəf	ssaxfəf	<i>make light</i>	10
		ssiβrəç	ssaβrəç	<i>make black</i>	6
		ssizwəʁ	ssaʒwəʁ	<i>make red</i>	7
		ssihləw	ssahləw	<i>make sweet</i>	4
		ssimləl	ssamləl	<i>make white</i>	5
		ssizəð	ssaʒəð	<i>make sweet</i>	1

c) Q-roots in a regular derivation. Finally, some Q-roots can form both ‘normal’ and Q-verbs. In these pairs, exemplified in (7), the corresponding ‘normal’ verb never contains any full vowel:

(7)	Q-verb	‘normal’ verb	
	iwsif	wsəf	<i>be broad</i>
	iʁxis	ʁxəs	<i>be cheap</i>
	iðjiq	ðjəq	<i>be narrow</i>
	iwsir	wsər	<i>be old</i>
	ismið	sməð	<i>be cold</i>

We conclude from these observations that Q-verbs vowels do not belong to the root. They are morphological markers in the Q-paradigm.

Consonant Gemination in the Perfective of Basic Q-verbs

In addition to apophony (cf. above), the perfective of certain Q-verbs displays a gemination of the second root consonant. (Or, the first root consonant in the case of the two verbs of the shape VC₁VC₂, aḏaj ‘be heavy’ and iziḏ ‘be sweet’.) (8a) and (8b) show that gemination cuts across verb classes that do and do not apophonyze in the perfective. (8c) shows that there seems to be some variation with respect to the realization of gemination: in some cases, the speakers produce both a form with and without gemination in the perfective. No gemination is observed in the perfective of derived causatives.

(8)	a. + gemination, — apophony	<i>aor.</i>	<i>pf.</i>	
		iβriç	βæriç	<i>be black</i>
		imzi(j)	mæzzi	<i>be small</i>
		iɣ ^w zif	ɣ ^w ædd ^z if	<i>be long</i>
		iwzil	wædd ^z il	<i>be short</i>
		izḏij	zæddij	<i>be clean</i>
		imlih	mællih	<i>be salty</i>
		imriɣ	mærrɣ	<i>be too salty</i>
		aḏaj	ḏḏaj	<i>be heavy</i>
b. + gemination, + apophony		ilwiɣ	lægg ^w aɣ	<i>be smooth</i>
		izwiɣ	zægg ^w aɣ	<i>be red</i>
		izfir	zæffar	<i>be tasteless</i>
		iwrɣ	wærraɣ	<i>be yellow</i>
		imlul	mællul	<i>be white</i>
		imsus	mæssus	<i>be tasteless</i>
		ismum	sæmmum	<i>be sour</i>
		ifsus	fæssus	<i>be light</i>
		ilqiq	lqaq / læqqaq	<i>be slack</i>
		isliw	slaw / sællaw	<i>be wilted</i>
c. +/- gemination, + apophony				

Gemination of the second root consonant can also be found in ‘normal’ verbs, but never in the perfective: in ‘normal’ verbs, gemination is a property of the intensive stem.

Adjective Formation

Q-verbs stand out with respect to adjective formation. Two strategies can be observed: i) the majority of adjectives require a specific suffix, *-an*, b) a restricted set of adjectives involves templatic formation.

Recall from sections “[Vocalic Properties](#)” and “[Consonant Gemination in the Perfective of Basic Q-verbs](#)” that the perfective of Q-verbs is characterized by two parameters: gemination of R₂ and apophony. Taking into account the values of these parameters, we can establish a descriptive generalization concerning the choice between *-an* suffixation and templatic formation: it appears that templatic adjective formation is restricted to Q-roots that exhibit both gemination and apophony in the perfective of the basic Q-verb. In all other cases, *-an* suffixation applies.

(9)	<i>Q-verb pf</i>		<i>adjective</i>
	— gemination	— apophony + apophony	a-...-an
	+ gemination	— apophony + apophony	
			a-CəCCaC

(10)	<i>Q-verb aorist</i>	<i>Q-verb perfective</i>	<i>adjective</i>	
	ihnin	ḥnin	aħninan	<i>with a good-heart</i>
	irzij	ṛzaj	aṛzajan	<i>sour</i>
	iβriç	β̣arriç	aβ̣ərçan	<i>black</i>
	izwiȳ	zəgg ^w ay	azəgg ^w ay	<i>red</i>

Normal (non-Q) verbs derive their adjectives either via templatic formation (11a) or via the prefixation of *am-* (which is also used to derive agentive nouns), or *an-* (11b). Some examples of numerically important classes of adjectives corresponding to normal verbs are given in (11).

(11)	<i>adjective formation</i>	<i>verb</i>	<i>adjective</i>	
a.	uCCiC	β̣qəs	uβ̣qis	<i>shattered</i>
	aCəCCaC	ħməq	aħəmmaq	<i>hasty</i>
b.	<i>am-CCaC-u</i>	zwir	aməzwaru	<i>first</i>
	<i>an-CaCC-u</i>	g ^w ri	anəgaru	<i>last</i>

Q-verbs never make use of the prefix *am/an-*. ‘Normal’ verbs never make use of the suffix *-an*. (There are a couple of exceptions with verbs that are morphologically regular, but express a quality, e.g., *fuh* ‘be dirty’, *afuħan* ‘dirty’.)

The Morphosyntactic Properties of the Perfective Stem of Q-verbs

The morphosyntactic properties of Q-verbs in Berber have been, and still are, the object of much debate in Berber linguistics (cf., among many others, Basset 1952; Galand 1955, 1980, 1990; Akouaou 1976; Chaker 1978; Bendjaballah 2006; Kossmann 2009; Allati 2015). In particular, the status of the perfective stem is still not well understood. In this section, we briefly review its three main

morphosyntactic properties: a reduced agreement paradigm, a particular semantic interpretation, and its behaviour when in combination with the directional particle *-d*.

Agreement in the Perfective

The first striking property of Q-verbs is that they display a reduced agreement paradigm in the perfective. As can be seen in (12), a normal verb like *xðəm* ‘work’ bears the prefix *j-* in the 3 ms.¹ By contrast, a Q-verb like *ixfif* ‘be light’ does not bear any prefix in the 3 ms: there is no overt subject agreement marker in this form.

(12)	<i>perfective.3ms</i>	a.	jə- 3MS	xðəm work.PF	“He worked.”	normal verb
		b.		xfif be light.PF	“He is light.”	Q-verb

In fact, the perfective paradigm of Q-verbs is characterized first by the absence of inflectional prefixes and second by the fact that there is a single inflectional marker in the plural: person and gender are not marked in the plural. This is shown in (13).

(13)	a. normal verb, perfective			b. Q-verb, perfective		
	1s		-ɣ	1s		-ɣ
	2s	θ-	-ð	2s		-ð
	3ms	i-		3ms		
	3fs	θ-		3fs		-θ
	1p	n-		1p		
	2mp	θ-	-m	2mp		
	2fp	θ-	-mt	2fp		-iθ
	3mp		-n	3mp		
	3fp		-nt	3fp		

This observation is true of the perfective paradigm only: in all other paradigms, Q-verbs do bear the normal subject agreement markers. This can be seen in (14), where the Q-verb *ixfif*, illustrated above in (12b), bears the 3ms /j/- prefix in the aorist and in the intensive. In these paradigms, Q-verbs behave on a par with normal verbs.

¹ The 3ms agreement marker surfaces in three different shapes: [j]-, [jə]- and [i]-. These realizations are determined by the phonotactic context (cf. Bendjaballah 1999 for an analysis of this allomorphy in a Government Phonology framework):

$$\begin{array}{rcl}
 /j/- & \rightarrow & [j]- \quad / \quad _V \\
 & & [jə]- \quad / \quad _CC \\
 & & [i]- \quad / \quad _CV
 \end{array}$$

(14)	a.	<i>aor.3ms</i>	<i>að</i>	<i>jə-_{3MS}</i>	<i>xðəm</i>	“He will work.”	‘normal’ verb
					<i>work.AOR</i>		
			<i>að</i>	<i>j-</i>	<i>ixfif</i>	“He will become light.”	Q-verb
				<i>3MS</i>	<i>light.AOR</i>		
b.	<i>ipf.3ms</i>			<i>i-</i>	<i>xəddəm</i>	“He is working.”	‘normal’ verb
				<i>3MS</i>	<i>work.IPF</i>		
				<i>jə-</i>	<i>tt^sixfif</i>	“He is becoming light.”	Q-verb
				<i>3MS</i>	<i>light.IPF</i>		

The Interpretation of the Perfective

The perfective paradigm stands out in another respect: it has a particular interpretation. While the imperfective forms of the verbs denote a process or a change of state, the perfective does not denote the perfective of that process or change of state (i.e., an event that has come to an end) as one would expect from the interpretation of the perfective of ‘normal’ verbs. Rather, the perfective of Q-verbs loses the eventive interpretation of the base verbs and denotes a property or pure state (“l’état pour lui-même, sans référence à un procès”, Galand 1980: 352).²

(15)	<i>βərriç</i>	“He is black.”
	<i>be black.PF</i>	
	<i>zəgg^way</i>	“He is red.”
	<i>be red.PF</i>	

The interpretation of the perfective of Q-verbs seems to be identical with that of adjectives used with the copula *ð*.

(16)	a.	<i>βərriç</i>	<i>uqərruj-is</i>	“His head is black.”
		<i>be black.PF</i>	<i>head.CS-POSS3s</i>	
		<i>aqərruj-is</i>	<i>ð aβərçan</i>	“His head is black.”
		<i>head.FS-POSS3s</i>	<i>COP black.MS</i>	

²As with normal verbs, the perfective stem is not marked for tense (i). Tense is marked on an auxiliary (ii):

i.	<i>assa</i>	<i>lqaj</i>	“Today it is deep.”
	<i>today</i>	<i>be deep.PF</i>	
	<i>iðəlli</i>	<i>lqaj</i>	“Yesterday it was deep.”
ii.	<i>yesterday</i>	<i>be deep.PF</i>	
	<i>jə-lla</i>	<i>lqaj</i>	“It was deep.”
	<i>3MS-be.PF</i>	<i>be deep.PF</i>	

- | | | | |
|----|--|---|--|
| b. | xfif-iθ
be light.PF-PL
t ^s ixfifanin
COP.F light.FP | | “They (f) are light.”
“They (f) are light.” |
| c. | jə-lla
3MS-be.PF
jə-lla
3MS-be.PF | lqaj
be deep.PF
ð alqajan
COP deep.MS | “It was deep.”
“It was deep.” |
| d. | a-tt ^s -af-əð
FUT-2S-find.AOR-2S
a-tt ^s -af-əð
FUT-2S-find.AOR-2S | əlqaj
be deep.PF
ð alqajan
COP deep.MS | “You will find that it is deep.”
“You will find that it is deep.” |

These semantic properties are not specific to particular lexical entries (viz. the roots of Q-verbs). Rather they characterize the Q-paradigm. Consider the roots that appear both in the Q-paradigm and in the regular one. A comparison of the semantics of the Q-verb with that of the regular verb establishes that only Q-verbs have the interpretation of a pure state. As illustrated in (17) with the root *əjq* ‘be narrow’, the perfective of the Q-verb expresses a pure state with no reference to a previous state whereas the perfective of the corresponding regular verb expresses a current state as opposed to a previous state. Modification of the semantics and modification of the paradigm go hand in hand.

- (17) a. *əjjijq*
be.tight.PF.3MS
“The trousers are tight.”
- b. *jə-əjjəq*
3MS-be.tight.PF
“The trousers are tight (now).”
(The person who wears it put on weight, or the trousers shrank in hot water)
- usərwal-ənni
trousers.CS-DEM
- usərwal-ənni
trousers.CS-DEM

The Effect of Directional –d

As far as we know, an important observation concerning the distribution of reduced agreement and the perfective stem of Q-verbs has not been made yet in the literature: Amazigh Bedar (pc) points out to us that the presence of the directional particle –d changes both the semantics and the agreement morphology of the perfective of Q-verbs.

Consider first the aorist and the intensive stem. In this context, the directional particle –d introduces telicity. It turns an atelic process into a telic one: (18a) may not have an endpoint, and (18b) must have one.

(18) a.	<i>aor.3ms</i>	<i>að</i>		<i>j-</i>	<i>izwiɣ</i>		“He will be reddening.”
		FUT		3MS	be red.PF		
	<i>ipf.3ms</i>			<i>jə-</i>	<i>tɛʔizwiɣ</i>		“He is reddening.”
				3MS	be red.IPF		
b.	<i>aor.3ms</i>	<i>a</i>	-d	<i>j-</i>	<i>izwiɣ</i>		“He will become red.”
		FUT	DIR	3MS	be red.PF		
	<i>ipf.3ms</i>			<i>jə-</i>	<i>tɛʔizwiɣ</i>	-əd	“He is becoming red.”
				3MS	be red.IPF	DIR	

For the perfective stem which, as we just saw, has a pure state reading, we might expect the absence of directional *-d*: states have no natural endpoint or result. However, directional *-d* does combine with the perfective of Q-verbs. When it does, it triggers important semantic and morphological effects. Semantically, it turns the pure state into a resultant state. In other words, a perfective Q-verb with *-d* refers to an event. (Note that the interpretation of the corresponding regular verb opposes two states in a discrete way, cf. (17b) while the interpretation of the Q-verb with *-d* in (19b) describes an event leading to a final state, as a continuum.) Morphologically, the construction is obligatorily marked with the full, regular, subject agreement paradigm, including the prefixes. In sum, the directional particle appears to coerce the perfective of the Q-verb into a regular verbal pattern (agreement and eventive semantics).

(19) a.		<i>ðəjjiq</i>		<i>usərwal-ənni</i>	“The trousers are tight.”
		be.tight.PF.3MS		trousers.CS-DEM	
		<i>ħnin</i>	-əθ		“She is sweet.”
		be sweet.PF	3FS		
b.	<i>i-</i>	<i>ðəjjiq</i>	-əd	<i>usərwal-ənni</i>	“The trousers shrank.”
	3MS	be tight.PF	DIR	trousers.CS-DEM	(*The person who wears it put on weight)
	<i>θə-</i>	<i>ħnin</i>	-əd		“She has become sweet.”
	3FS	be sweet.PF	DIR		

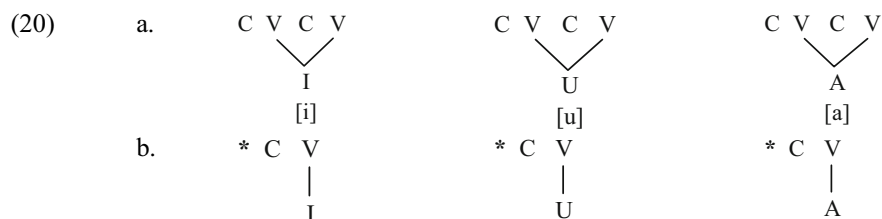
The Formal Properties of Verbal and Nominal Templates

Apparently evident decisions on the classification of data may have non-trivial consequences for the analysis. More than anywhere else, this truism is worth being reminded in a paper on Berber Q-verbs. As a matter of fact, there is a longstanding debate on the nominal origin of the perfective stem of Q-verbs at the diachronic level in Berber linguistics (for details and hypotheses, cf. e.g., Galand (1980, 1990), Kossmann (2009), Allati (2015)). The following quote by Galand (1980: 358) is representative of an influent position in this debate: “Le problème posé en berbère [...] est d’accueillir dans son système verbal un syntagme nominal qui, devenu forme conjuguée, n’en garde pas moins la valeur d’un statif.” Taking this position

seriously in a formal synchronic study, forces us not only to compare the template of the perfective of Q-verbs with the template of nouns: it might lead us to conclude that what appeared to be the “paradigm of Q-verbs” is really not one paradigm, but a disjunction of several (verbal and nominal) paradigms, the cells of which cannot be related to one another in any immediate way.

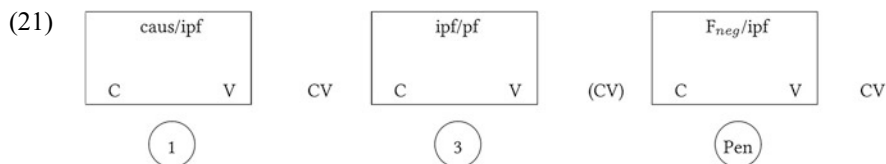
In this section, we explore the consequences of a nominal analysis of Q-verbs, building on earlier work on the internal plural of nouns. We first introduce the ‘normal’ (non-Q) verbal template and the template of internal nominal plurals, respectively. Then, we pass on to the Q-template of perfective and non-perfective forms, respectively and observe that it shares all of its active positions with the nominal, rather than the verbal template. We conclude with a brief exploration of the empirical and theoretical consequences of this observation.

Before we proceed, recall that segmental length alternations are pervasively accompanied by melodic alternations in Berber languages. Following earlier work (Lowenstamm 1991; Bendjaballah 1999), we assume that this includes vowels. In particular, we assume that a phonetically realized full vowel in Taqbaylit is phonologically long, i.e., it occupies the V-positions of two subsequent CV units. If an underlying full vowel fails to have access to two subsequent V-positions, it is de-linked from the skeleton and phonetically realized as either *schwa*, or zero depending on its right consonantal context.



The ‘Normal’ Verbal Template

The Taqbaylit Berber ‘normal’ (i.e., non-Q) verb template has the shape given in (21) (Bendjaballah 2007). It includes three infixal positions separated by root-positions: the first position, the third position, and the penultimate position. In the schematic illustration below, affixal CV units are boxed.



These affixal positions host markers of several morpho-syntactic categories. Crucially, any given position is not associated with a fixed morpho-syntactic value. Those values vary depending on the value of all other affixal positions in the template. Let us briefly review them – starting from the left periphery of the template.

Position [1] hosts either the causative or the imperfective marker. The causative is the prefix *s(s)-*, which surfaces in the forms in (1b). The imperfective marker is the prefix *t(t)^s-*, which surfaces in the second column of (1a.iii, iv, v, vi). Note that the causative and the imperfective prefixes are in complementary distribution: there is not a single causative verb that marks its imperfective via prefixation (second column of (1b)). This justifies the assumption of a single position for both prefixes.

Position [3] always hosts an aspectual marker, either the perfective or the imperfective. The perfective is marked by apophony, e.g., *a* → *u* in (1a.vi) and the vowel which is the target of apophony is associated with position [3]. In (1a.i) and (1a.ii), the imperfective is marked by the gemination of the second root consonant, which we assume takes place on position [3], too.

Finally, the penultimate position [Pen] hosts the imperfective aspect or another category, which we label F_{neg} . F_{neg} is a category selected by the negation: more specifically, it is activated under the negation in the perfective. The activation of [Pen] can be seen in the very last column of (1a). Take as an example the verb in (1a.i): in the negative perfective, an [i] surfaces immediately before the last root consonant, which was absent in the positive perfective. [Pen] also hosts the imperfective marker in the derived causatives. This becomes apparent when considering the forms in the second column of (1b): a vowel ([a] or [u]) surfaces on the very same site as the [i] of the negative perfective immediately before the last root consonant.

The first important observation concerning the template in (21) is that there is no 1-to-1 correspondence between the templatic positions and the categories they realize. Except for [3], which consistently marks an aspectual category, the affixal positions in the verbal template can realize quite different features: [1] hosts the causative and the imperfective, and [Pen] hosts the imperfective and a category selected by the negation.

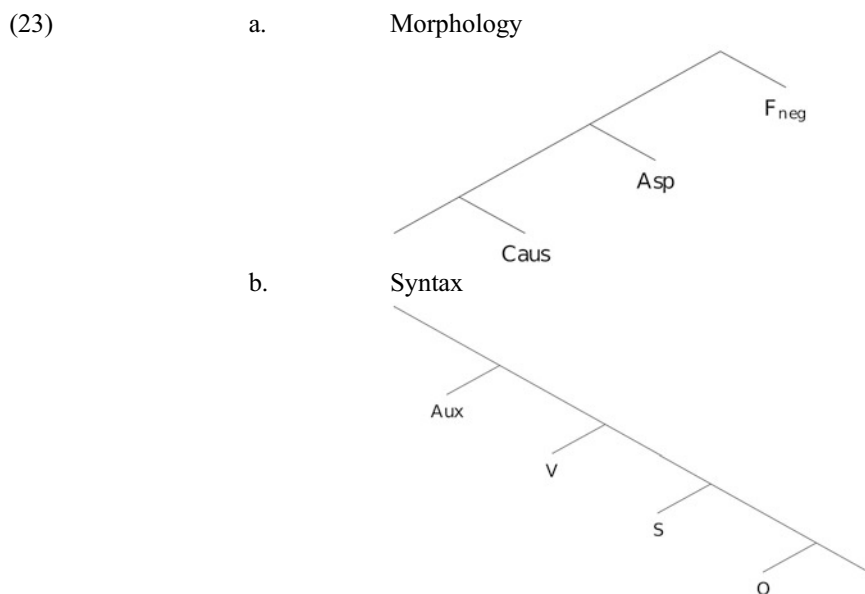
(22) Attested configurations

	[1]	[3]	[Pen]												
a. one active position	<table> <tr><td><i>CAUS</i></td></tr> <tr><td><i>IPF</i></td></tr> <tr><td></td></tr> <tr><td></td></tr> </table>	<i>CAUS</i>	<i>IPF</i>			<table> <tr><td></td></tr> <tr><td></td></tr> <tr><td><i>IPF</i></td></tr> <tr><td><i>PF</i></td></tr> </table>			<i>IPF</i>	<i>PF</i>	<table> <tr><td></td></tr> <tr><td></td></tr> <tr><td></td></tr> <tr><td></td></tr> </table>				
<i>CAUS</i>															
<i>IPF</i>															
<i>IPF</i>															
<i>PF</i>															
b. two active positions	<table> <tr><td><i>CAUS</i></td></tr> <tr><td><i>CAUS</i></td></tr> <tr><td></td></tr> </table>	<i>CAUS</i>	<i>CAUS</i>		<table> <tr><td></td></tr> <tr><td><i>PF</i></td></tr> <tr><td><i>PF</i></td></tr> </table>		<i>PF</i>	<i>PF</i>	<table> <tr><td><i>IPF</i></td></tr> <tr><td></td></tr> <tr><td><i>F_{NEG}</i></td></tr> </table>	<i>IPF</i>		<i>F_{NEG}</i>			
<i>CAUS</i>															
<i>CAUS</i>															
<i>PF</i>															
<i>PF</i>															
<i>IPF</i>															
<i>F_{NEG}</i>															

It is worth noting that the simultaneous activation of all three templatic positions appears to be excluded in the verbal domain. In particular, the perfective under

negation of the causative does not exhibit penultimate lengthening, as we would expect it to happen from its non-causative counterpart.

Bendjaballah (2012, 2014) observes that the arrangement of attested markers in the verbal template is related to the syntactic c-command. First, the lexical category “causative” is always the left-hand member of any pair (22b). Second, the category selected by the negation is always the right-hand member of any pair. In between, we find the aspectual categories. A clear generalization then emerges: for any pair of infixes, the one located to the right must be syntactically higher than the one located to the left. In other words, the linear order in the template is head-final (23a). In addition, note that Berber clause structure is head-initial: in syntax, heads precede their phrasal complements (23b). Bendjaballah (2014) concludes from this that the order of the markers in the verb template obeys the Mirror Principle in its linear interpretation: any given templatic marker in the verbal template c-commands its left-hand neighbour.



The Nominal Template

The main classes of Taqbaylit nouns are illustrated in (24). The first column gives masculine nouns – the second column gives their feminine counterpart. (24a) provides singular forms and (24b) provides the corresponding plural forms. Note that Berber nouns can be divided into three groups depending on their plural formation: nouns with an external plural (e.g., *axxam* (sg), *ixxam-ən* (pl) ‘house’), nouns with an internal plural (e.g., *amfiʃ* (sg), *imfaʃ* ‘cat’), and nouns with a mixed plural

(e.g., *ajəttum* (sg), *ij^wəðm-a-n* (pl) ‘shoot’). The relevant facts concern internal and mixed plurals. For our purposes here, it is sufficient to illustrate internal plurals.

(24) a. *Singular*

Masculine Feminine

VCəCCVC θ- VCəCCVC -θ aβərnus ‘burnous’ θaβərnusθ ‘burnous dim.’

VCCVC θ- VCCVC -θ amfɪʃ ‘cat’ θamfɪʃθ ‘she-cat’

VCVCVC θ- VCVCVC -θ ajazið ‘rooster’ θajazið ‘hen’

b. *Internal Plural*

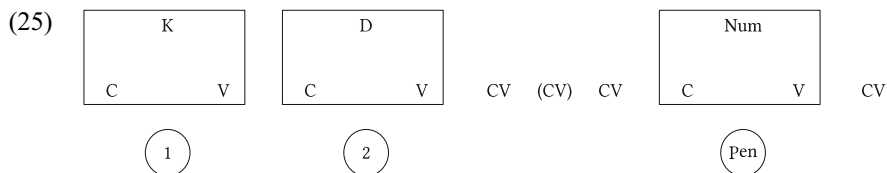
Masculine Feminine

VCəCCVC θ- VCəCCVC -θ iβərnas ‘burnous’ θiβərnas ‘burnous dim.’

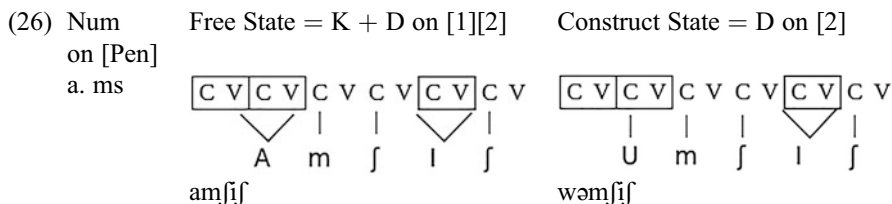
VCCVC θ- VCCVC -θ imfɪʃ ‘cat’ θimfɪʃ ‘she-cat’

VCVCVC θ- VCVCVC -θ ijuzað ‘rooster’ θijuzað ‘hen’

Bendjaballah and Haiden (2008) argue that Taqbaylit nouns have the template in (25). (On the basis of a detailed examination of Taqbaylit nominal plural formation, Ben Si Said (2014) concludes that the plural is characterized by a CVCVCVCVCV template with the penultimate position identified by A in the internal plurals, e.g., θimfɪʃ ‘she-cats’. This result concurs with the template shown in (25).) The left periphery of the noun hosts two templatic sites. The initial/outer position expresses K, the second/inner position hosts D. Turning now to the right periphery, Bendjaballah and Haiden (2008) identify gender, class, and number markers in different positions. For the purposes of the present paper, it suffices to consider a single position, the penultimate, which hosts a marker related to number (for details and discussion, cf. Bendjaballah and Haiden 2008, 2013; Bendjaballah 2014).



This template is illustrated with noun type VCCVC in (26) and (27) (masculine and feminine, respectively). The first column shows the forms in the Free State and the second column the forms in the Construct State. In the Construct State, [1] remains empty. (The emptiness of the initial site in the CS is correlated with the specific syntax of this state, cf. Bendjaballah and Haiden (2013) for details.)



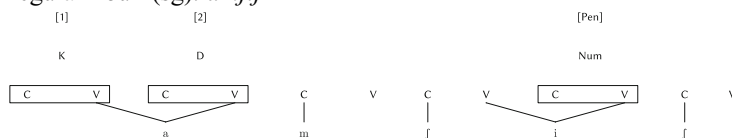
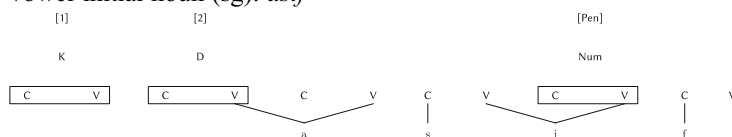
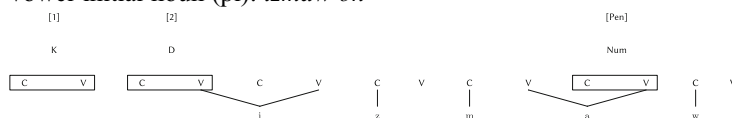
- b. mp
- | | |
|-------|--------|
| | |
| imfaʃ | jəmfaʃ |
- (27) Num on [Pen]
- a. fs
- | | |
|---------|---------|
| | |
| θamʃiʃθ | θəmʃiʃθ |
- b. fp
- | | |
|--------|--------|
| | |
| θimfaʃ | θəmfaʃ |

Three immediate observations must be made about the nominal template: first, the two affixal positions at the left edge of the template are not separated by a root CV-position. This contrasts with the situation observed in the verbal template. Second, assuming with Guerssel (1987, 1992), and Bendjaballah and Haiden (2008, 2013) that K is structurally higher than D, we observe a head-initial configuration among the active positions of the nominal template. Third, there is no limit as to the simultaneous activation of affixal positions in the nominal template: as illustrated with the FS of feminine internal plurals in (27b), all three positions can be simultaneously activated. Again, this contrasts with the situation observed in the verbal template. An examination of the more complex nominal patterns reveals an asymmetry between the left edge (position [1] and [2]) and the rest of the nominal template. While the left edge is head-initial, the right edge (including, but not limited to Pen) appears to be head-final. This asymmetry corresponds with the extended syntactic domains of N and D (DP for the left edge, NP for the right edge), and the position of post-nominal demonstratives and prenominal numerals, respectively.

One nominal class appears to be particularly relevant for our purposes due to the defective nature of its left periphery, vowel-initial nouns. In this class (28b), contrary to regular nouns (28a), the initial full vowel of the Free State is preserved in the Construct State:

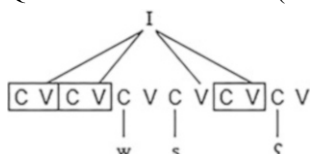
- | | | | | | |
|------|---------------------------|------------|-----------------|--------|----------|
| (28) | | Free State | Construct State | | |
| a. | <i>Regular noun</i> | sg | amʃiʃ | wəmʃiʃ | ‘cat’ |
| b. | <i>Vowel-initial noun</i> | sg | asif | wasif | ‘stream’ |

Basset (1952) and Guerssel (1983) – among others – analyzed the difference between these two classes as follows: the initial vowel in the Free State of regular nouns (the one that disappears in the Construct State) is an affix; the initial vowel in the Free State of ‘vowel-initial’ nouns is part of the root.

(31) a. regular noun (sg): *amfiʃ*b. vowel-initial noun (sg): *asif*c. vowel-initial noun (pl): *izmaw-ən*

The Template of Q-verbs

With this in mind, consider now the template of Q-verbs. The aorist of a representative Q-verb is represented in (32). Q-verbs are characterized by three infixal positions that correspond to the positions occupied by the vocalic melody. These positions have morpho-syntactic value and appear as boxed CV in the representations. The presence of three affixal positions in the template is common to nouns, ‘normal’ verbs, and Q-verbs. ‘Normal’ verbs differ from Q-verbs and nouns in the arrangement and the activation patterns of the three infixal positions.

(32) Q-verb: *iwsif* ‘be broad’ (aor.)

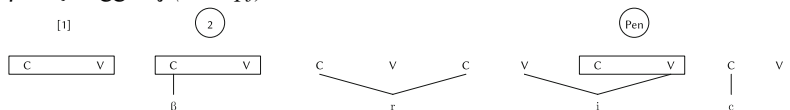
In the template of Q-verbs and nouns, the two affixal positions at the left edge are adjacent. They occupy the first and second position of the template (25, 32). In the template of ‘normal’ verbs, the two affixal positions at the left edge are separated by a root position (21). This supports a comparison of Q-verbs with nouns and lends credibility to the historical scenario: As far as the linear order of affixal positions is concerned, Q-verbs clearly pattern with nouns, rather than with verbs.

Consider now the set of the configurations attested in the paradigm of Q-verbs:

(33) a. *wsis/hlaw* (*base.pf*)



βərriç/zəgg^way (*base.pf*)



b. *ssiwsəs* (*caus.aor*)



ssawsəs (*caus.pf*)



c. *iwsis* (*base.aor*)



tt^siwsis (*base.int*)

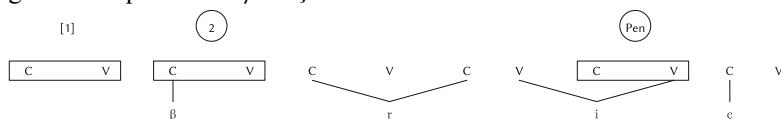


ssiwsis (*caus.int*)



The crucial configurations are given in (33c): in both intensive forms (the causative and the base) and in the aorist of the base all three templatic sites are simultaneously activated. In particular, we do not observe the reduction of the vowel in penultimate position that we had found in the perfective of the causative under negation in ‘normal’ verbs. In other words, the constraint on maximal realization, typical of the ‘normal’ verbal template, does not apply in the paradigm of Q-verbs. This is another important feature Q-verbs share with nouns.

Now let us turn to the perfective forms in (33a, 34). Both the systematic emptiness of the external position [1] and the spreading of root material into the inner position [2] in the class with gemination suggest a comparison with vowel-initial nouns.

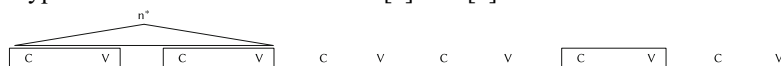
(34) a. perfective without gemination: *wsif*b. geminated perfective: *βarriç*

However, contrary to what we observe in the perfective of Q-verbs, the initial affixal positions of vowel-initial nouns can host independent markers: inflectional material appears in position [2] and light prepositions in position [1] (Bendjaballah and Haiden 2013). The perfective of Q-verbs does not tolerate any marker at its left edge. Not even template-external, subject-agreement prefixes are present. Why should this be so?

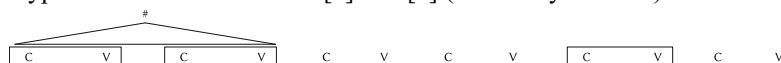
Two major options can be envisaged concerning the absence of prefixes in the perfective of Q-verbs. First, we can strictly apply the parallelism with nouns and hypothesize that positions [1] and [2] actually realize silent non-verbal markers in the perfective of Q-verbs. This first scenario (35a) considers that verbal prefixes cannot attach to the left of a perfective Q-stem because positions [1] and [2] realize nominal features excluded by the subcategorization frame of verbal agreement prefixes. This first hypothesis comes close to treating the perfective of Q-verbs as a sort of participle. It cannot explain without further assumptions (e.g., silent auxiliaries) why such elements can nevertheless head finite clauses.

The second hypothesis (35b) would still have positions [1] and [2] present at the templatic level. However, according to this second hypothesis, these positions host no affixal morpheme at all. Assume that, in analogy with pauses and boundary tones, such empty sites are interpreted in syntax as a boundary marker (or a ‘Grenzsignal’ in Trubetzkoy’s terms). The boundary marker is equally excluded by the subcategorization frame of verbal agreement prefixes. Given furthermore that Q-verbs normally realize aspectual and argument structure markers in positions [1] and [2], as illustrated in (33), it seems reasonable to expect that the emptiness of these positions in the (underived) perfective entails the loss of eventive semantics.

(35) a. Hypothesis 1: nominal markers on [1] and [2]

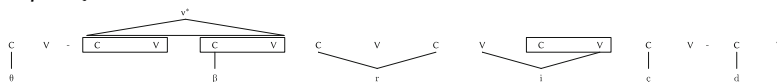


b. Hypothesis 2: no markers on [1] and [2] (‘boundary marker’)



Both hypotheses are compatible with the effect of directional *-d* for our informant. Its presence coerces the projection of a semantically eventive, syntactically verbal category at the left edge of a Q-verb, even in the perfective. The presence of this marker triggers the affixation of (regular) agreement prefixes.

(36) $\theta\text{-}\beta\text{ərri}\check{\text{ç}}\text{-}\text{əd}$



Conclusion and Outlook

We conclude from these observations that the template of Q-verbs does bear important resemblances with the template of nouns. In particular, the prosodic topography of Q-verbs bears out affixal CV units at the same sites as nouns: two adjacent initial positions and the penultimate position are templatically active in Q-verbs. As in the template of nouns, all three of these positions can simultaneously host affixal markers. In the template of regular verbs, the two initial positions at the left edge are separated by a root position and a complexity constraint applies: no more than two internal positions can be simultaneously active. Taken together, we conclude that Q-verbs are prosodically nominal.

In the final part of the chapter, we have tried to link certain key features of the perfective of Q-verbs to the nominal pattern (but not necessarily nominal morpho-syntax) of their template. We compared two hypotheses. The first one, which relies on a strong link between the prosodic form of templates and the markers hosted by the respective positions pushes the synchronic reality of diachrony to its limit. It analyses the perfective of Q-verbs as a nominal form. The credibility of this hypothesis suffers from the fact that it cannot explain without further assumptions why such nominal forms can head finite clauses.

A second hypothesis crucially relies on the independence of prosodic patterns (i.e., templates) and lexical items (i.e., morphemes hosted by templatically active positions). If templates can exist as prosodic patterns independently of the nature of the markers they host, then it is envisageable that Q-verbs are nominal by their prosodic pattern only, but verbal or defectively verbal in morpho-syntax. As this second hypothesis can do without additional ad hoc assumptions like empty auxiliaries in the underived perfective, it appears to be more credible than the first one.

Let us finally point to a prediction resulting from our analysis of directional *-d*. We propose that *-d* coerces a verbal-eventive analysis of the initial, empty positions of the perfective of Q-verbs. As an operation on lexical entries, coercion is predicted to exhibit significant inter-speaker variation. In particular, conservative speakers who reject coercion should likewise reject directional *-d* in the perfective of Q-verbs.

References

- Akouaou A (1976) L'expression de la qualité en berbère: le verbe. Thèse de 3^e cycle, Ecole Pratique des Hautes Etudes. [Impr. 2013. L'expression verbale de la qualité en amazighe. Rabat: Institut Royal de la Culture Amazighe]
- Allati A (2015) Les verbes d'état berbères: des fossiles linguistiques revisités. *Studi Afr Quaderni Studi Berberi Libico-berberi* 4:87–116
- Bader Y (1984) Kabyle Berber phonology and morphology: outstanding issues. PhD, University of Illinois
- Bader Y (1985) Schwa in Berber: a nonlinear analysis. *Lingua* 67(2–3):225–249
- Basset A (1952) *La langue berbère*. Oxford University Press, London
- Ben Si Said S (2014) De la nature de la variation diatopique en kabyle: étude de la formation des singulier et pluriel nominaux. Thèse de doctorat, Univ. de Nice
- Bendjaballah S (1999) Trois figures de la structure interne des gabarits. Thèse de doctorat, Univ. Paris Diderot
- Bendjaballah S (2001) The negative Preterite in Kabyle Berber. *Folia Linguistica* XXXIV (3–4):185–223
- Bendjaballah S (2006) Les “verbes de qualité” en kabyle. *Etudes et documents berbères* 24:71–102
- Bendjaballah S (2007) Constraints and markers in templates. In: Picchi M-C, Pona A (eds) *Proceedings of the XXXII Incontro di Grammatica Generativa*. Edizioni dell'Orso, Alexandria, pp 21–340
- Bendjaballah S (2011) Note sur les noms à voyelle initiale stable en kabyle. In: Mettouchi A (ed) *Parcours berbères, Mélanges offerts à Paulette Galand-Pernet et Lionel Galand pour leur 90^e anniversaire*. Berber studies, vol 33. Rüdiger Köppe, Cologne, pp 417–434
- Bendjaballah S (2012) *La grammaire des gabarits*. Habilitation à Diriger des Recherches. Université Paris Diderot, Paris
- Bendjaballah S (2014) Remarks on nonconcatenative affixation. *BAALL* 6:45–73
- Bendjaballah S, Haiden M (2008) A typology of emptiness in templates. In: Hartmann J, Hegedus V, van Riemsdijk H (eds) *Sounds of silence*. North Holland linguistic series, linguistic variations, vol 63. Elsevier, Amsterdam, pp 21–57
- Bendjaballah S, Haiden M (2013) The representational anomalies of floating markers: light prepositions in Taqbaylit of Chemini. In: Biberauer T, Roberts I (eds) *Challenges to linearization*. *Studies in generative grammar*, vol 114. Mouton, Berlin, pp 331–375
- Chaker S (1978) *Un parler berbère d'Algérie (Kabyle)*. Publications de l'Université d'Aix-en-Provence, Aix-en-Provence
- Galand L (1955) Etat et procès: les verbes de qualité en berbère, vol XLII. *Hespéris*, pp 245–251
- Galand L (1980) Une intégration laborieuse: les ‘verbes de qualité’ du berbère. *Bull Soc Ling Paris* LXXV(1):347–362
- Galand L (1990) Du nom au verbe d'état, le témoignage du berbère. In: Mukarovsky HG (ed) *Proceedings of the fifth international Hamito-Semitic Congress, 1987, Hamito-Semitic, Chadic, Berber*. *Beiträge zur Afrikanistik* 40, vol 1. Universität Wien, Wien, pp 123–138
- Guerssel M (1983) A phonological analysis of the Construct State in Berber. *Ling Anal* 11:309–330
- Guerssel M (1987) The status of the lexical category preposition in Berber: implications for the nature of the Construct State. In: Hale K (ed) *Studies in Berber syntax*. Center for Cognitive Science, MIT Press, Cambridge, MA, pp 159–190
- Guerssel M (1992) On the case system of Berber. *Can J Linguist* 37:175–195
- Guerssel M (1996) Ablaut in classical Arabic measure I active verbal forms. In: Lecarme J, Lowenstamm J, Shlonsky U (eds) *Studies in Afroasiatic grammar*. HAG, The Hague, pp 123–134
- Guerssel M, Lowenstamm J (1996) The derivational morphology of the Classical Arabic verbal system. Ms. UQAM and Univ. Paris 7

- Idrissi A (2001) Towards a root-and-template approach to shape-invariant morphology. PhD, UQAM
- Kossmann MG (2009) La flexion du prétérit d'état en berbère: questions de morphologie comparée. In: Chaker S, Mettouchi A, Philippson G (eds) Etudes de phonétique et linguistique berbères. Hommage à Naïma Louali (1961–2005). Peeters, Paris, pp 155–176
- Lowenstamm J (1991) Vocalic length and syllable structure in Semitic. In: Kaye AS (ed) Semitic studies in Honor of Wolf Leslau on the occasion of his 85th birthday. Harrassowitz, Wiesbaden, pp 949–965

The Status of Berber (Past and Present) and Its Contact with Arabic

18

Moha Ennaji

Contents

Introduction	386
A Historical Background	386
The Contact Between Berber and Colloquial Dialectal Arabic	388
Colloquial Dialectal Arabic as an Indigenous Language	388
Colloquial Arabic Loans	390
The Issue of the Alphabet	393
Introducing Amazigh in Schools	396
The Evolution of the Status of Berber	398
Conclusion	400
References	400

Abstract

This chapter deals with the status of Berber (Amazigh for natives) (I am using the term “Berber” interchangeably with the word “Amazigh”. The natives prefer the latter because it has positive overtones (Amazigh means free man; Tamazight, free woman; Imazighen, free people), whereas the term Berber is considered derogatory because it stems from the Roman term “Barbari” (Barbarian) referring to the people that the Romans had conquered. However, the term “Berber” is so commonly used by researchers today that it has become neutral. The status of Berber has changed from a spoken non-recognized language to a written official language in 2011 in Morocco and in 2014 in Algeria thanks to the militancy of the Amazigh cultural movement and the political will to integrate Amazigh into development. Berber has been in contact with Arabic since the seventh century, and the impact of Arabic on it is mirrored in the quantity of borrowings and loanwords that have entered Berber. Reciprocally, Berber has also influenced mainly Colloquial Arabic at the lexical and the morphological levels. The

M. Ennaji (✉)

Sidi Mohamed Ben Abdellah at Fez, International Institute for Languages and Cultures, Fez, Morocco

teaching of Berber, which has been officially launched in both countries since 2003 and which has been steadily growing, will contribute to language change, to the standardization and codification of the language, and will consolidate its new status as an official language.

Introduction

This chapter is divided into four sections. Section one deals with the features of Berber and Arabic linguistic contacts with a focus on borrowings. Section two is concerned with the political and linguistic issue of the Berber script. Section three reflects the changes that have occurred since the introduction of Berber in schools and universities in the Maghreb. Finally, section four deals with the extent to which the new status of Berber has favourably changed attitudes toward it and toward its revival.

A Historical Background

Before discussing the issue of language contact between two indigenous languages, Berber and Colloquial Dialectal Arabic, it is necessary to provide a historical and a sociolinguistic background.

Berber is an Afro-Asiatic language that is mainly used north of the Great Sahara and in North-West Africa. It is the mother tongue of the first inhabitants of North Africa. It is spoken in Morocco, Algeria, Tunisia, Libya, Egypt, Mauritania, Mali, Niger, Nigeria, and Chad.

Four major population groups may be distinguished. The largest population group is in Morocco, where the Berber-speaking population totals about 15 million people, i.e., 40% of the population. The second largest group is represented by Algeria, where more than 7 million people speak Berber (25% of the population). Third, there are the Tuareg populations in the sub-Saharan countries, namely Mali, Niger, and Libya, which are estimated to be about 1 million people. Fourth, there exist approximately 140,000 Berber-speaking people scattered in isolated areas in Siwa (Egypt), Tunisia, and Mauritania. The regions where Berber is spoken are discontinuous, as they are usually surrounded by populations speaking other languages, such as Arabic (Ennaji 2005, 2019). According to many scholars, there are nine major varieties of Berber: (1) Tashelhit, spoken in southern Morocco, (2) Tamazight in central Morocco, (3) Tarifit in northern Morocco, (4) Kabyle in Tizi-Ouzou (Algeria), (5) Mzab in Ghardai (Algeria), (6) Shawiya in Aures (Algeria), (7) Tuareg in Southern Algeria, Niger, and Mali, (8) Tamashek in Niger, Mali, and Nigeria, and (9) Tamahaq in Libya and Nigeria. However, it is in Morocco that Berber is vastly used by the population. See Ennaji (2019), Imarazene (2010), Boukous (2004). In Tunisia, Berber is spoken in the extreme south (in the region of Medenine). There are less than six Berber villages where the

inhabitants speak Berber, located in Djerba, west of Matmata and east of Gafsa. (Cf. Payne 1983). After the “Arab” Spring in 2011, there is a growing interest in the revival of the language and in the sociolinguistics of Tunisia. See Gabssi (2011) and Daoud (2011).

These dialects belong to the same Berber language although they tend to be mutually unintelligible, especially when the two varieties in question are in extreme geographical points; for example, a speaker of Tarifit in the north of Morocco cannot understand nor communicate with a Tashelhit speaker from the south. The lack of mutual intelligibility between most of these varieties may be ascribed chiefly to the lack of a standard form of the language, which leads to its quasi exclusion from the public life and its murky introduction in the education system (see Chaker 1984; Ennaji 2009). Geographical isolation and absence of the language in the media can be cited as other major causes of unintelligibility.

Berber speakers resort mainly to Arabic for writing (Brahimi 2000; Benrabah 2014). However, despite the Arabization policy, the assimilation of Berbers has been slow and was recently rejected by the Berber cultural movement because of the rising consciousness about Berber identity which demands to promote it.

However, this traditional division of Berber according to rural areas does not accurately reflect the effect of the huge rural exodus that the Maghreb has experienced since independence. Consequently, there are now important Berber populations in the big cities of the Maghreb and Europe such as Casablanca, Rabat, Agadir, Algiers, Utrecht, Paris, Brussels, and Madrid (see Chaker 1984).

In addition, the exact number of the varieties is not clearly specified, since there is only the linguistic atlas of the Rif by Lafkioui (2007), and Behnstedt and Woidich (2010) atlas on Arabic varieties.

On the other hand, many linguists consider Berber essentially a single language, despite its many varieties (Chaker 1984).

Bilingual Berbers often speak Colloquial Arabic as well. This is particularly true of the younger generation and those less than 50 years old. Monolingual Berbers can generally be found among the older generation of illiterate people (mostly women) and preschool children living in the mountainous regions. The various Berber dialects in the Maghreb are the vehicle of traditional oral literature. Berber enjoys a rich oral heritage of songs, legends, poems, proverbs, and short stories. France, Belgium, Holland, and Germany host a large community of North African immigrants from Berber origin, who have been struggling to preserve their native language and culture. The Berber language was revived specifically in France, where a Berber Academy was founded in 1967, and where several cultural activities are organized (cf. LeSaout 2012).

Berber plays a secondary role in the mass media. A Berber radio channel broadcasts its programs daily in Morocco and Algeria, but not in Tunisia. In Morocco, the Berber radio is on air from eight in the morning until midnight, with each dialect having its period of broadcasting per day. A Berber TV channel (*attamina*) was launched in March 2010. TV news is broadcast daily—since August 1994—on the government owned TV channel 2 M once a day. In Algeria, there is a

radio station which broadcasts exclusively in Berber, and there is also a Berberophone newscast, in addition to an all-Berber TV channel.

After years of struggle by the Berber movement activists, Berber was finally recognized for the first time in history as an official language in the Moroccan constitution of 2011, alongside Modern Standard Arabic (Ennaji 2019; Benrabah 2014). The Algerian constitution of 2002 declared Berber a national language. In 2014 it was declared official. In both countries, Berber is no longer considered a dialect. However, although Berber is codified, as it is now written in the Tifinagh alphabet in Morocco and in Latin or Tifinagh in Algeria, it is still not completely standardized (see Ennaji 2009). In spite of the recent changes in the status of Berber and efforts to promote it at all levels, the domains in which it is used are still largely restricted to home and informal transactions. However, Berber is a vital and living language, and there is a strong feeling among Berber native speakers and the progressive forces in the region that, to be preserved, Berber ought to be codified, standardized, and used in education, media, and public administration (see Sadiqi 1997; Ennaji 1997, 2009).

In the following section, we will consider the linguistic contacts between two indigenous languages: Berber and Colloquial Dialectal Arabic.

The Contact Between Berber and Colloquial Dialectal Arabic

Colloquial Dialectal Arabic as an Indigenous Language

Ferguson (1972) states that the Arab-speaking world, in general, is characterized by *diglossia* in the sense that two varieties of Arabic coexist, namely Modern Standard Arabic (or “high” Arabic) and Colloquial Arabic (or the “low” Arabic variety). The first is codified, standardized, associated with the Holy Qur’an, and represents a great literary tradition. Colloquial Arabic is the language of everyday conversation and is neither codified nor standardized, hence its low status. In North Africa, Berber and Colloquial Dialectal Arabic predate French colonization, as they have coexisted for centuries.

Like the Amazigh language, Colloquial Arabic is a spoken language widely used by the Moroccan speech community. More than 60% of Berbers are bilingual, as they speak Colloquial Dialectal Arabic as a second language. The latter indigenous language is the mother tongue of more than half of the population. Colloquial Dialectal Arabic is almost completely a spoken language, although some informal letters, plays, or texts may be written in this variety using the Arabic script. It is referred to as *darija*, or *šammiya*, i.e., the language of the common people.

Berberophones use Colloquial Arabic as a lingua franca when they meet Berber speakers from different geographical areas. It is spoken in informal settings (at home, in the street, and with friends and relatives) by educated and uneducated people in their everyday speech.

Colloquial Arabic can be divided into urban/Sedentary (mdini) and rural (Bedouin or froubi) Arabic. However the sociolinguistic situation is a bit more complex in large cities, namely Casablanca where Colloquial Arabic is considered Bedouin. Equally, the tendency in various Arabic-speaking regions is for Bedouin dialects to override Sedentary ones (Ennaji 2013).

Until recently, there was not much contact between speakers of the rural variety and those of the urban variety. But today, with the help of transportation facilities and the relative industrialization of the cities, many rural people have taken to the cities in search of a better life. As a result, rural Moroccan Arabic speakers can be encountered even in large cities.

There are also religious, political, and educational reasons that make Modern Standard Arabic and Colloquial Arabic two different varieties of Arabic (Ennaji 2005: 47–70). Moroccans (and Arabs on the whole) consider Modern Standard Arabic a prestigious language, and the only form worth learning in schools, while they see Colloquial Dialectal Arabic as a corrupt and vulgar dialect (Ferguson (1972) [1959]). As mentioned earlier, Modern Standard Arabic is generally associated with literacy and learnedness, while Colloquial Dialectal Arabic is associated with illiteracy and orality.

However, the existence of Colloquial Arabic as a vital mother tongue is a reminder that Moroccans have a rich oral literature and cultural tradition. Colloquial Arabic is stigmatized, chiefly because it is associated with illiteracy. A person who speaks only Colloquial Arabic is considered illiterate. Historically and in religious contexts, writing has been the exclusive domain of Modern Standard Arabic; for this reason, the average Moroccan would not write in his own mother tongue. Although Colloquial Arabic is one of the essential distinguishing characteristics of Moroccan culture and the mother tongue of more than half of the population, it is excluded from education, official settings, and the printed media, as mentioned earlier.

Colloquial Arabic may expand and become more widespread in the future, given its growing place in the mass media. Many Berber-speakers adopt Colloquial Dialectal Arabic when they move from rural to urban areas because it is the language of transactions in urban areas. Radio and television, at times, use Colloquial Arabic alternatively with Modern Standard Arabic which will, in the long run, reduce the gap between speakers of Colloquial and Modern Standard Arabic. The adoption of Colloquial Arabic by the media contributes to its spread across the country.

Although Colloquial Arabic is in competition with other languages and varieties, it occupies, like Berber, an important position as an oral language. It is the mother tongue of the majority of the population, and in this capacity it reflects cultural authenticity. Hence, it is essential to preserve both indigenous languages, Colloquial Arabic and Berber, and reinforce them in formal and informal domains of use (Ennaji 2002).

Nonetheless, Berber has a lower status than Colloquial Arabic among the Arabic-speaking population and has a significant influence on Colloquial Arabic but no impact on Modern Standard Arabic, the formal written variety. Since Berber is passed on from generation to generation by the oral medium, Its speakers resort to Arabic for writing (Brahimi 2000; Benrabah 2014).

Both Berber and Colloquial Arabic are under pressure by the economic, political, and cultural forces of globalization and modernization, which seek to marginalize minority languages. In the worldwide surge of wealth and information technology, the currency of local languages, like Berber and Colloquial Arabic, is devalued.

In the following section, we will describe and analyze the extent to which Colloquial Arabic has influenced Berber language, and how this process has led to language change and shift.

Colloquial Arabic Loans

Berber has always entertained contacts with other languages such as Arabic for historical reasons. Arabic has been influencing Berber since the Arab conquest of North Africa in the seventh century. Similarly, French has had a strong impact on Berber; especially at the lexical level since the French invasion of North Africa in 1830. As a result, Berber is full of French loans (e.g., *fsyan* ‘officer’, *lkufa* ‘convoy’, *ttira* ‘field’, *tabla* ‘table’), etc. (cf. Ennaji 1991).

However, it is undeniable that Arabic loans are the most common, because Arabic is the language of prestige in the Maghreb, and it is the language of the Qur’an, the holy book of North African Muslims, Arabic-speaking, and Berber-speaking alike. In the current situation of the Amazigh lexicography, it is difficult to determine the extent of Arabic borrowings and loans, but according to Chaker (1984), Arabic loans represent 40% in the Berber lexicon in Morocco, and 25% of Kabyle lexicon, in Algeria.

Many linguistic accounts of the Berber language concentrate on its interaction with Arabic (see Taifi 1979). Despite this language contact, Berber speakers continue to use their own language along with Arabic as a second language, a common linguistic phenomenon often called an *adstratal* language, whereby the shift is incomplete and the speakers continue speaking their native language along with the dominant language Arabic for our purposes. The situation can be qualified as one involving a superstratum, where Modern Standard Arabic may be considered “superstrate”. At any rate, all these concepts are related to language contact, irrespective of whether they cause language shift or not.

Substratal language occurs when speakers of a language shift from their own language to use another language as a result of language contact. (An adstratum or adstrate is a language that is in contact with another language in a neighbor population without having identifiably higher or lower prestige. The concepts ‘adstrate’ and ‘substrate’ are related to language contact in general, whether leading to language shift or not). This is the case of extinct languages like Coptic in Egypt, which disappeared because of the strong dominance of Arabic. Linguistic research has often underscored the impact of *substratal* and *adstratal* languages on the development of Berber and Arabic, particularly, their role in the emergence of the

new forms of Arabic that are, nowadays, known as the Arabic dialects (Versteegh 2001).

In their approach to language contact, Thomason and Kaufman (1988) propose a model for linguistic change as the outcome of language contact. They argue against a strictly intra-linguistic account of language change. Even though they do not assert that all linguistic changes have to be accounted for by outer factors, i.e., by language contact, they do emphasize that internal factors are less significant than is assumed by most linguists (see Versteegh 2001).

Thomason and Kaufman (1988: 21) distinguish between two kinds of contact-induced change, which they called “borrowing”, and “substratum interference”. In the former type, foreign elements are integrated into the speaker’s native language—in the latter type, items from the speaker’s native language are integrated into the foreign language. In their model, then, ‘borrowing’ as the use of words or elements from a foreign language into the native language, has the same meaning as in other frameworks (see also Kossmann 2017; Bensoukas and Boudlal 2012; Youssi 1992; Caubet 1990).

Thomason and Kaufman use the notion of ‘substratum interference’, however, in a different way from standard research. They argue that interference is part and parcel of language shift, as speakers of a language A are in the course of espousing language B and discarding their own language. But during the acquisition process, their language is deficient, because they make mistakes which are in turn introduced by the learners into their new language. This leads to language change or shift, which may gradually affect the language of the whole community. Nevertheless, Thomason and Kaufman’s model does not seem to apply to Berber-Arabic language contact. It could be utilized to argue that Colloquial Arabic developed as a result of some Berber speakers shifting to Arabic. Perhaps that is one of the reasons why not many Berber loans can be found in the Maghreb Arabic varieties, while the morphosyntactic influence is quite evident (Bensoukas and Ech-Charfi 2018; Ennaji 1985).

It should be noted that contacts between Amazigh and Arabic led to new words and phrases whose lexical origin is not always easy to specify. The influence of Arabic is very strong, especially in the religious vocabulary, e.g., *rəbbi*, ‘God’, *lixert* ‘the beyond’, *ddin* ‘religion’, *nbi* ‘prophet’, *lislam* ‘Islam’, *lqiyyama* ‘Judgment Day’, *tažallit* ‘prayer’, *zzka* ‘alms’, *luđu* ‘ablution’, *lhidj* ‘the pilgrimage’, *lihsan* ‘charity’, etc. The lexicon of commerce and politics is also heavily influenced by Arabic. The Amazigh language has borrowed many Arabic words related to trade: e.g., *ttesbiq* ‘advance’, *ssalaf* ‘credit’, *ttaman* ‘price’, *lbi f u šra* ‘trade’, etc.

There are also Arabic words borrowed from the field of education, administration and policy (see Taifi 1979): e.g., *lmedrasa* ‘school’, *lustad* ‘professor’, *limtiħan* ‘exam’, *lintixab* ‘elections’, *lmurrššah* ‘candidate’, *lbərlaman* ‘parliament’, *lqanun* ‘law’, *lbaladiya* ‘municipality’, *nnayb* ‘member of parliament’, *rrays* ‘president’, *ljama3a* ‘commune’, *lhizb* ‘party’, *lħala lmadaniya* ‘marital status’, *lwizara* ‘ministry’, *lhukuma* ‘government’, *lmuħami* ‘lawyer’, *lkatib l3am* ‘general secretary’,

ddimuqratiTiya ‘democracy’, *ttanmiya* ‘development’, *lmisTara* ‘rule / procedure’, *lqayd* ‘caid’, *lhakm* ‘leader’, *lqadi* ‘judge’, *listinatif* ‘appeal’, etc.

Arabic borrowings are generally adapted to the morpho-phonology and syntax of Berber. Phonetic changes allow Arabic borrowings to integrate more easily into the vocabulary of the Berber language: e.g., Berber *akafir* of Arabic *kafir* ‘ungodly’; *ayzzar* for *agzzar*, or *gzzar*, ‘butcher’ in Arabic; Berber *axddar* for the Arabic *xddar* ‘greengrocer’; and *abnnay*, borrowed from Arabic *bnnay* ‘mason’. The integration of Arabic loanwords into Amazigh is facilitated by the genealogical relationship of the two languages, both of which belong to the Hamito-Semitic family, nowadays called Afro-Asiatic.

But despite the possibility of integrating Arabic borrowings into the Berber language, the strong presence of Arabic has pushed some Arabic-Berber bilinguals to borrow the original form from Arabic, avoiding forms obtained by derivational transformation, as in the following examples; *lisr* ‘left’, *limn* ‘right’, *rriyada* ‘sport’, *lkulliya* ‘faculty’, *lwadifa* ‘function’, etc. Both Colloquial and Modern Standard Arabic loans are the result of Berber-Arabic bilingualism, which has kept growing since the implementation of the Arabization policy in the 1960s. The number of these bilinguals has increased since Colloquial Arabic is spoken even in rural areas where the Berber language was, hitherto, the only spoken language.

On another level, Berber has also influenced Colloquial Dialectal Arabic, mainly at the lexical and morphological levels. Many words of Berber origin have entered Colloquial Arabic as in the following examples: *seksu* ‘couscous’, *atay* ‘tea’, *berra* ‘outside’, *sarut* ‘key’. Morphologically, Colloquial Moroccan Arabic has a derivational noun circumfix /*ta*-. .-*t*/, which is borrowed from Berber. For illustration, we have *ta-xddar-t* ‘activity of a greengrocer’, *t-agzzar-t* ‘being a butcher’, *t-anjaar-t* ‘being a carpenter’, *t-abnnay-t* ‘act of building’, etc. Most of these Arabic words with the Berber circumfix indicate a profession or a social activity. This circumfix is highly productive on native Arabic noun stems but uncommon on borrowed Berber stems. This pattern of productivity is taken to be evidence in support of morphological borrowing (cf. Steinkruger and Seifart 2009) and challenges the principle of a “borrowability hierarchy” (cf. Haugen 1950) in which lexical elements are borrowed before grammatical morphemes. Moreover, the prefixal particle of the Arabic circumfix, /*ta*-, is a complex form of the Berber /*t*-/ feminine + /*a*-/ absolute state; for example, *tazzla* ‘running’, *taRuri* ‘studying’, *tamgra* ‘harvesting’, etc. But this prefix is not a productive borrowing in Colloquial Dialectal Arabic. The borrowed circumfix is, however, a case of borrowing which does not conform to traditional typologies of language change.

From this discussion, and within the framework of Thomason and Kaufman (1988: 38–39), I assume that language contact triggers language change when loanwords and new items are introduced into the language, irrespective of whether

these items belong to the speech of the new learners or to that of the native speakers. Thomason and Kaufman also refer to a process of adoption of the new elements by the speech community at large. This is the case of Arabic loans which are incorporated in Berber, where the language change and the loans have spread to the whole speech community (see Versteegh 2001). It seems that what Thomason and Kaufman say about the social factors controlling linguistic borrowing is not well exploited in relation to Arabic-Berber contact.

Thomason and Kaufman's model correlates with the standard studies of borrowing, such as Haugen (1950) and Weinreich (1968 [1953]), who do not distinguish rigorously between the two types of change, and who argue that borrowing can go from phonological through morphological to syntactic borrowing. In this view, syntactic borrowing always presupposes morphological and phonological borrowings, which in their turn presuppose lexical borrowing (cf. also Romaine 1989: 50–66). However, the spread of both lexical and non-lexical borrowing depends on the degree of integration: a phenomenon that needs a further discussion in another research.

It is notable that the degree of lexical borrowing from Arabic into Berber is, by far, more substantial than the reverse borrowing from Berber into Arabic. Many Arabic loans have entered Berber directly through the influence of Arabic-Berber bilinguals, rather than via Moroccan or Algerian Arabic monolinguals (Owen 2000).

Given the recent official recognition of Berber in Morocco and Algeria, its introduction in schools is expected to gradually lead to its codification and standardization and play a great role in literacy and basic education in North Africa (see Ennaji 2019; Benrabah 2014). In this case, Berber is an interesting tool for preserving and disseminating popular culture and orality in the region. To attain this goal, the Tifinagh alphabet has been selected to write Berber, as we shall discuss in the section below.

The Issue of the Alphabet

The Tifinagh alphabet, which is over 2000 years old, has been officially adopted as the alphabet of Berber since 2003 in Morocco. It is widely referred to as Tifinagh-Ircam. See the table below. Tifinagh has been sporadically employed in some parts of North Africa in the writing of Amazigh; however, the most commonly utilized alphabet is Latin. As a case in point, the writings of the well-known Kabyle writer, the late Mouloud Mammeri (1917–1989), and those of contemporary Kabyle writers are in Latin script, supported by diacritics and phonetic symbols.

	Tifinagh	Latin Equivalent	Arabic Equivalent	Examples
1	ⵜ	b	ب	ⵜⵜⵓⵔⵉⵏ (road)
2	ⵎ	m	م	ⵎⵓⵏⵓⵔ (ground)
3	ⴼ	f	ف	ⴼⵓⵏⵓⵔ (hand)
4	ⵜ	t	ت	ⵜⵓⵔⵓⵔⵓⵔ (knowledge)
5	ⵏ	d	د	ⴼⵓⵏⵓⵔⵏ (knee)
6	ⵉ	<u>I</u>	ط	ⵜⵉⵉ (eye)
7	ⵉ	<u>D</u>	ض	ⵉⵉⵓ (foot)
8	ⵓ	s	س	ⵓⵔⵓⵔ (tongue)
9	ⴰ	z	ز	ⵉⵏⵓⵔⵓⵔⵓⵔ (Amazigh)
10	ⵔ	<u>S</u>	ص	ⵔⵉⵉ (blow)
11	ⴰ	<u>Z</u>	ژ	ⵉⵔⵓⵔ (bile)
12	ⵏ	n	ن	ⵉⵓⵏⵓⵔ (wheat)
13	ⵏ	l	ل	ⵉⵏⵓⵔⵓⵔⵓⵔ (white)
14	ⵓ	r	ر	ⵓⵓⵓⵔ (play)
15	ⵓ	<u>R</u>	ر..	ⵜⵓⵔⵓⵔⵓⵔ (outside)
16	ⵉ	C	ش	ⵉⵉⵓⵓⵓⵔ (crowd)
17	ⵏ	j	ج	ⵉⵏⵓⵔⵓⵔⵓⵔ (bald)
18	ⵏ	k	ك	ⵜⵉⵔⵓⵔⵓⵔ (pot)
19	ⵏ	k°	ك	ⵉⵏⵓⵔⵓⵔⵓⵔⵓⵔ (friend)
20	ⵏ	g	گ	ⵏⵓⵔⵓⵔⵓⵔ (my borther)
21	ⵏ	g°	گ	ⵉⵏⵓⵔⵓⵔⵓⵔⵓⵔ (red)
22	ⵏ	x	خ	ⵜⵉⵔⵓⵔⵓⵔ (goat)
23	ⵓ	q	ق	ⵉⵓⵓⵓⵔ (bag)
24	ⵓ	γ	غ	ⵉⵓⵓⵓⵔ (bread)
25	ⵏ	h	ح	ⵉⵏⵓⵔⵓⵔⵓⵔ (dance)
26	ⵏ	ε	ع	ⵉⵓⵓⵓⵔ (cloth)
27	ⵓ	h	هـ	ⵉⵓⵓⵓⵔⵓⵔ (blanket)
28	ⵏ	y	ي	ⵉⵓⵓⵓⵔⵓⵔ (horse)
29	ⵓ	w	و	ⵉⵓⵓⵓⵔ (water)
30	ⵓ	a	ا	ⵉⵓⵓⵓⵔ (mountain)
31	ⵏ	i	ي	ⵉⵓⵓⵔ (mouth)
32	ⵓ	u	و	ⵉⵓⵓⵔ (face)
33	ⵓ	e		ⵉⵓⵓⵔⵓⵔⵓⵔ (I read)

The Tifinagh script has been efficiently modernized by the Royal Institute of Amazigh Culture (IRCAM) in Morocco and has been officially recognized by the World Association of Alphabets. It is now used in textbooks and websites, including government ones, TV, public signs, and in many teaching programs that are developed by IRCAM in collaboration with the Ministry of Education

(cf. Boukous 2004; Ameur and Boumalk 2004, 2006; El Moujahid 2006; Agnaou 2008). IRCAM, created by a royal decree in October 2001, was established exclusively to promote the Amazigh language and culture. The IRCAM's publications and endeavors to standardize the Amazigh language and introduce it in schools will have a positive impact on its teaching in Morocco as well as on its future in all of North Africa.

In Algeria, some schools use Latin, while others adopt Tifinagh, as has been shown by Dadoua Hadria and Boumedini (2020). In Morocco, a few Berber creative writers use Tifinagh. Others use the Latin alphabet or the Arabic script. The autumn of 2002 saw a hot media debate over the alphabet issue. Proponents of the Latin alphabet justified their attitude by claiming that it was practical and close to the phonetic system, which itself derives from the Roman alphabet. Many Berber non-government organizations (NGOs) argue that the Latin script is better suited for writing Berber because it can be used on the Internet, and in word processing on the computer. They also prefer the Latin alphabet to the Arabic one, which lacks vocalization (Imarazene 2010). Those who prefer the Arabic alphabet claim that it is the closest to the Berber language roots and because, historically, both Arabic and Berber belong to the Chamito-semitic/Afro-asiatic family. Proponents of the Arabic script, especially Islamists, argue that since most Berbers are Muslims who read Arabic, they will accept writing/reading Berber in Arabic script more readily than in Latin script (see the Islamist newspaper *Attajdid* of October 25, 2002).

To avoid the conflict between Islamist fundamentalists and Berber activists, Tifinagh was officially chosen in 2002 as a medium solution. This decision was based on a recommendation by IRCAM that was reached by a majority vote at its January 30–31, 2002 session held in Rabat. Initially, IRCAM presented pros and cons of three scripts: Tifinagh, Arabic, and Latin. For some activists, the choice was meant to settle the conflict between Islamists and Berberists, but this is debatable (cf. Sadiqi 2014; Maddy-Weitzman 2011; Asid 2000).

This decision has been welcomed by the majority of Berberophones and Arabophones as a good political middle solution to the dilemma of which script to use for Berber (see Ennaji 2003). In general, for many Berber native-speakers, Tifinagh is a good choice because it strengthens Berber identity, consolidates the language autonomy, and shows that Berber culture is one of the oldest in the region, as it goes back more than 2000 years. For others, although Tifinagh is politically a correct decision in the present context as it has prevented a confrontation between the Islamists and the Berber activists, it is pedagogically impractical because it has the huge drawback of being an obscure system and a third script for Moroccans to learn. This group argues that it is useless to codify Berber in a script that people do not know and may find hard to grasp and in which literature and written materials are lacking (see Bentolila 2003). In fact, most Berber literature is oral; however, a few literary texts are available in both Arabic and Latin scripts (Imarazene 2010; Errihani 2008).

To avoid the considerable difficulty that might be encountered in reading and comprehending materials written in Berber (with the Tifinagh alphabet), an adequate

orthography must be developed to ensure the smooth introduction of Berber in the school system and to gain the support of the people and of the elite (Sadiqi 2014).

Introducing Amazigh in Schools

Little research has been undertaken to examine the learning achievement of Berber children in school and to see the impact of second language instruction on the literacy skills and on the basic school subjects. Penchoen (1968), who studied the difference in learning achievement between Arabic-speaking and Berber-speaking children in Tunisia, found that there was no real discrepancy between the two groups. This result is due to the high motivation of Berberophone children to learn Arabic. This may also be due to the fact that Modern Standard Arabic is like a second language to Colloquial Arabic native speakers. Similarly, in Wagner (1993: 175), Berber children had, globally, the same ability to read Arabic as their Arabic-speaking classmates. Wagner (1993: 176) states that “there appears to be some advantage to speaking dialectal Arabic as a mother tongue when first beginning to read, but any advantage diminishes substantially over subsequent years of schooling”.

As far as reading achievements are concerned, Wagner (1993: 178) observes that Berber and Arabic-speaking children’s performance is the same. Berberophone and Arabophone children who attended Qur’anic (Islamic) preschools did better in reading achievement than those who did not go to preschool. However, urban children who were sent to private preschools outperformed their non-preschooled or Qur’anic preschooled counterparts in literacy.

Children whose mother tongue is Colloquial Arabic slightly outperform Berber-speaking children who had no preschooling experience, especially in the first years of school. This is due to the linguistic similarity between Colloquial and Standard Arabic, and the possibility of positive transfer from the former to the latter. But after a few years of learning, both groups have similar competence in Arabic reading and writing. Berberophone children catch up with their Arabic-speaking counterparts after four to five years of progress toward Berber-Colloquial Arabic bilingualism. The more the Berberophone children’s competence in Colloquial Arabic increases, the more their Standard Arabic learning achievements progress.

In Morocco, the Amazigh language and culture were introduced in 300 primary schools in September 2003 and in 5431 schools in 2010, all across Morocco in rural and urban areas (source: Direction for Statistics and Planning, Ministry of Education, 2010). The series of textbooks *Tifawin* and *Anlmd Tamazight* are used in most of these schools. Teachers of Amazigh are committed to doing the job effectively and to carry out the teaching of the language.

In Algeria, the Amazigh language is taught, not only in primary schools, as in Morocco, but also in secondary schools, using either the Tifinagh Alphabet or the Latin script. However, neither in Morocco nor in Algeria, is Amazigh used as a vehicle of instruction in schools (Imarazene 2010).

To carry out the teaching of Amazigh in an effective way, linguistic and pedagogical training is both urgent and necessary. The training has to take into account the linguistic competence of the teachers and the pedagogical methods and techniques. Continuous training will help the teachers to update their pedagogical knowledge and competence, which will improve the quality of teaching the Amazigh language and culture.

Unlike the teachers of Arabic and French who have been to teacher training schools (e.g., *Ecole Normale Supérieure*), those of Amazigh have benefited from only a few training sessions and workshops. It is high time that Amazigh teachers had the same training opportunities. IRCAM strives to ensure continuous training and retraining. However, the time allotted to teacher training is unsatisfactory. The sessions are infrequent and seldom focus on teachers' pedagogical needs.

In addition, the time allotted to the teaching of Amazigh language (three hours a week) is in the opinion of many teachers and activists insufficient, and in some schools only Tifnagh alphabet is taught and sometimes drawing lessons are taught instead of Berber due to a lack of teachers, or a for lack of motivation on the part of the administration or teachers.

The Tifnagh alphabet is often a problem after the second year in primary school, for it is difficult for students to read a long text in Tifnagh because the latter is not usually used on a daily basis in real life. Audiovisual aids are also lacking in most schools, for there are few multimedia tools to help teach Amazigh in a proper way, which ought to be part of the programs and methods used to teach the language (Ennaji 2003).

For comparison, the official teaching of Amazigh in Algeria was initially launched at the university level with the opening of a department of Amazigh language and culture at the University of Tizi-Ouzou in 1990 and a second department at the University of Bejaia in 1992. These departments first began with Masters programs to train Amazigh students in Amazigh linguistics, literature, and civilization. This training came to an end in 1997, when these departments changed direction by training undergraduates who would later become teachers of Amazigh in high schools.

Note that the teaching of Amazigh at the University of Tizi-Ouzou is conducted predominantly in French, while it is ensured, in large part, in Tamazight in both the Bejaia and Bouira University Centers, where a third department of Amazigh language and culture was launched in September 2008. The teaching of content courses is done in French for lack of terminology of a metalanguage in Amazigh (Imarazene 2010). Although the number of Amazigh students has risen in high schools and universities, the teaching of Amazigh is still, neither mandatory nor generalized for all schools and areas in either Algeria, or Morocco (Abrous 2010).

In Morocco, at the level of higher education, many universities have established Majors and Masters degrees in Amazigh studies, namely the Universities of Oujda, Fès, Agadir, and Tetuan, since 2007 and 2009. Thus, these Moroccan universities have engaged their human potential and logistics in the process of promoting Amazigh culture. However, this situation should not blind us to the great contributions of many Moroccan researchers who did their doctoral theses on Amazigh

language and literature, either in Morocco, Europe, or North America. Today, some of them indulge in the teaching of Amazigh studies at university level.

However, this experiment suffers from many problems such as shortage of references and publications on, and in, Amazigh language, as well as the scarcity of teachers. To overcome these obstacles, most Amazigh study programs seek the help of teaching assistants in order to teach a few courses (cf. Saa 2010).

In the following section, let us turn to discuss attitudes toward Berber language and culture.

The Evolution of the Status of Berber

The status of Berber has evolved over the years. In the seventies, it was considered merely a dialect, and attitudes of non-Berber speakers varied from indifference to negative; but since the 1990s, official attitudes have become favorable to Berber, thanks to the pressure from academics and civil society activists. To calm down the demonstrators in the city of Errachidia in the South East of the country who chanted slogans in favor of promoting Berber, the late king Hassan II took a politically important step by deciding, on August 20, 1994, to introduce Berber in primary schools as a first step. Later on, his successor, King Mohammed VI founded the IRCAM, whose goal was to codify, standardize, and promote Berber in all sectors, namely in education and media. In July 2011, the Amazigh language was recognized as an official language in Morocco. It has been recognized as a national language in Algeria since 2002. In 1995, the equivalent of IRCAM in Algeria “le Haut Commissariat à l’Amazighité” saw the light of day (Imarazene 2010).

Until the creation of the IRCAM in Morocco and the Haut Commissariat à l’Amazighité in Algeria, there were no official institutions or public funds allotted to encourage research on the Berber language and culture. Berber studies were carried out by Moroccan and non-Moroccan researchers and students on an individual noninstitutional basis. Thus, in Morocco alone, hundreds of BA monographs and dozens of MA and PhD theses were written on Berber. There exist also publications in, and on, Berber culture and literature (see Ennaji 2005: Chapter 4).

Most Moroccan and Algerian political parties have not taken a clear stand on the issue of Berber until very recently. Post-independence, the nationalist movement in both countries focused on Arabic and Arabization, neglecting the Berber language and culture. The major concern of the nationalists in post-independence was nation–state building and the national union. It was only in the 1990s that political parties started to break their silence and show some timid support to the Berber cultural demands.

The first Moroccan political party to have asked for the official recognition of the Berber language and culture as national components of the Moroccan culture was the “Mouvement Populaire”, a rural party led by a Berber charismatic politician, Mahjoubi Aherdan. The socialists were indifferent to Berber until the late King Hassan II announced in 1994 its introduction in schools. However, only the “Parti du Progrès et du Socialisme” publicly stated in its 1995 national convention that the

Berber language was a national language and a part and parcel of the Moroccan culture.

Similarly, in Algeria, only two political parties supported the Berber movement from the start: “Le Front des Forces Socialistes” (FFS), founded in 1963 by Hocine Ait Ahmed, and the Rally for Culture and Democracy (RCD), a secular political party associated with the “Berber Cultural Movement”, founded in February 1989 by Saïd Sadi, a human rights activist.

Overall, there is a national consensus in both Algeria and Morocco that the Berber language and culture are basic components of society that are to be promoted and preserved, for they express cultural authenticity and identity. The Berber language and culture can also be used to strengthen national solidarity and communication among the various segments of the Maghrebi population.

The demands of the Amazigh in Morocco are different from those in Algeria. In Morocco, the Amazigh question has integrative goals in the sense that Moroccan Amazighs demand their integration in national economic, social, and cultural development. Additionally, Morocco is characterized by cultural and ethnic diversity as there are Moroccans of Amazigh, Arab, African, or Andalusian origin. The fact that the present King’s mother is Amazigh and his father was Arab is significant. By contrast, in Algeria the Amazigh movement is political, as it seeks the autonomy of Berber regions, especially Kabylia: On 2 June 2010, the Algerian militant, Ferhat Mhenni, announced in Paris the creation of a provisional Kabyle government in exile called “gouvernement provisoire kabyle” (Cheref 2006) (See his declaration to the electronic journal DNA. URL: <http://www.dna-algerie.com/interieure/ferhat-mhenni-notre-objectif-est-l-autonomie-regionale-de-la-kabylie-2>. Accessed on 15 Oct. 2012).

Amazigh activists today use peaceful means to advocate their cause: they use art (songs, poetry, music), new communication technology (computers, Internet), and social media to help spread Amazigh and to join forces in the struggle for the promotion of Amazigh language and culture. Through dozens of websites and electronic journals like “Amazighworld” and “Tamazgha”, the Amazigh cause has gone global (see Cornwell and Atia 2012).

The use of the social media has given Berber a transnational platform (see Cornwell and Atia 2012). Prior to the Internet, the Amazigh identity was an internal question, in the sense that the Amazighs of Morocco did not know about other Amazighs in Algeria, Tunisia, Libya, or Mali. Given the fact that the Amazigh populations are in discontinuous zones and are divided regionally as subgroups, each assumed that their problems were local and did not have any significance to others.

Furthermore, a large number of Berberophone activists and intellectuals have committed their lives to establishing awareness of the Amazigh existence. Academics, linguists, researchers, and artists contributed to the revitalization of Berber language and culture. As a result of such a commitment, the authorities in Morocco and Algeria have been pressured to recognize the Berber language and culture. With the availability of satellite television, computers, the Internet, smartphones, YouTube, Facebook, WhatsApp, DVDs, and CD-Roms, the Berber population had the opportunity to articulate its distinct identity.

Berber activists seek to promote Amazigh as a tool for the democratization and laicization of the society and for sensitizing people of the advantages of multilingualism and multiculturalism. The youth have begun to accept their Amazigh identity and to reconcile with their origins and cultural authenticity (Asid 2000; Salhi 2000; Ennaji 2005: Chapter 4).

Conclusion

This sociolinguistics of Berber conforms to the general characterization of the interface between indigenous languages and their socio-cultural context. I have considered, albeit very briefly, a number of issues relevant to the interface between two indigenous languages in the Maghreb, Berber, and Colloquial Arabic, including the linguistic contacts and the impact of loans and borrowings upon them.

In light of this work, the following conclusion can be drawn regarding the evolution of the Berber language in the Maghreb: being regarded as a language and no longer as a set of dialects, Berber language has recently entered public space as it is no more limited to the confines of the home, close friends, and rural areas of Morocco and Algeria. It has been introduced into the education system and in the media, in both countries. The recent official recognition of Berber in Morocco and Algeria has offered a real opportunity for reflection and debate and mirrors the fact that officials' and non-officials' attitudes towards the Berber language have evolved over the recent two decades. The key, for now, is to work on the standardization of the Berber language in order to integrate it into the education system as a tool and object of study.

References

- Abrous N (2010) L'Enseignement du Tamazight en Algérie. *Langues et Linguistique* 25(26):11–28
- Agnaou F (2008) *Tamawalt inu Tamzwarut*. IRCAM, Rabat
- Ameur M, Boumalk A (eds) (2004) *Standardisation de l'amazighe*. Publication of IRCAM, Rabat
- Ameur M, Boumalk A (2006) (Dir.) *Structures morphologiques de l'amazighe*. Publication of IRCAM, Rabat, Série : Colloques et Séminaires No 10, Imprimeri Maarif Eljadida
- Asid A (2000) Berber in the discourse of political Islam (in Arabic). *Publications of Taousna Journal*, Casablanca
- Behnstedt P, Woidich M (2010) *Wortatlas der arabischen Dialekte: Band I: Mensch, Natur, Fauna und Flora*. Brill, Leiden
- Benrabah M (2014) Tensions between Arabophones and Berberophones in Algeria: historical development and current issues. In: Ennaji M (ed) *Multiculturalism and democracy in North Africa*. Routledg, London, pp 57–80
- Bensoukas K, Boudlal A (2012) An Amazigh substratum in Moroccan Arabic: the prosody of schwa. *Langues et Littératures* 22:179–221
- Bensoukas K, Ech-Charfi A (2018) Amazigh-Arabic language contact phenomena: introduction. *International journal of Arabic Linguistics* (Special Issue: Arabic-Amazigh Contact. Edited by K. Bensoukas and A. Ech-Charfi) 4(1):1–9
- Bentolila F (2003) Interview with *Le monde Amazigh*, November 2003, No 39

- Boukous A (2004) La standardization de l'amazighe: quelques prémisses . In: M Ameur, A Boumal (eds), pp 11–22
- Brahimi F (2000) Loanwords in Algerian Berber. In: Owen J (ed) Arabic as a minority language. Mouton de Gruyter, Berlin, pp 371–382
- Caubet D (1990) Arabe maghrébin: Passage à l'écrit et institutions. *Faits de Langues* 13:235–244
- Chaker S (1984) Textes en linguistique berbère. CNRS, Paris
- Cheref A (2006) Engendering or endangering politics in Algeria? *J Middle East Women's Stud* 2(2): 60–85
- Comwell G, Atia M (2012) Imaginative geographies of Amazigh activism. *Soc Cult Geogr* 13(3): 255–274
- Dadoua Hadria N, Boumedini B (2020) Des Acquis Considérables pour la Langue et La Culture Amazighes en Algérie: Etat des Lieux. *Langues et Linguistique* 44/45:17–26
- Daoud M (2011) The sociolinguistic situation in Tunisia: language rivalry or accommodation. *IJSL* 211:5–16
- El Moujahid E (2006) Morphologie du nom de qualité en amazighe. In: Ameur M, Boumalk A (eds) Structures Morphologiques de l'Amazighe. IRCAM, Rabat, pp 151–159
- Ennaji M (1985) Complex sentences in Arabic, Berber, and English: A contrastive analysis. Königshausen I Neumann, Würzburg
- Ennaji M (1991) Aspects of multilingualism in the Maghreb. *Int J Social LLang [Sociolinguistics of the Maghreb. Special Issue]* 87:7–25
- Ennaji M (1997) The sociology of Berber: change and continuity. *Int J Sociol Lang [Berber Sociolinguistics. Special Issue]* 123:23–40
- Ennaji M (2002) Language contact, Arabization policy and education in Morocco. In: Rouchdy A (ed) Language contact and language conflict in Arabic. Routledge/Curzon, London, pp 70–88
- Ennaji M (2003) Attitudes to Berber and Tifinagh. In: Proceedings of the conference on Amazigh language and culture. Al-Akawayn Univeristy, Ifrane, pp 16–23
- Ennaji M (2005) Multilingualism, cultural identity and education in Morocco. Springer, New York
- Ennaji M (2009) Réflexions sur la standardisation de la langue amazighe. *Revue Asinag* 2:5–16
- Ennaji M (2013) Multilingualism in Morocco and the linguistic features of the Casablanca variety. In: Aronin L, Fishman J, Singletont D, Laoire MÓ (eds) Current multilingualism. Mouton, Berlin, pp 231–246
- Ennaji, M. (2019). The Berber cultural movement in the Maghreb: contemporary issues in transnationalism. In A. de Toro and J. Tauchnitz, The world in movement: performative identities and diasporas (pp. 238–251). Leiden: Brill
- Errihani M (2008) Language policy in Morocco: implications of recognizing and teaching Berber. Lightning Source Incorporated, Chicago
- Ferguson, C. (1972)[1959] Diglossia. In PP Giglioli (ed), Language and social context (pp 232–251). Penguin, Harmondsworth
- Gabsi, Z (2011). Attrition and maintenance of the Berber language in Tunisia. *Int J Sociol Lang* 2011(211):135–164
- Haugen E (1950) The analysis of linguistic borrowings. *Language* 26(2):10–23
- Imarazene M (2010) Tamazight à l'Ecole Algérienne: Entre Progression et Régression. *Langues et Linguistique* 25(26):29–40
- Kossmann MG (2017) Berber-Arabic language cod-switching in Imouzzar du Kander (Morocco). *Lang Typol Universals* 65(4):369–382
- Lafkioui M (2007) Atlas linguistique des variétés berbères du Rif. Rüdiger Köppe Verlag, Cologne
- LeSaout D (2012) L'amazighité et l'immigration en France: Transformation des liens sociaux et productions d'identités. In: Ennaji M (ed) Migration et Hybridité. Publication Centre Sud Nord, Fès, pp 81–92
- Maddy-Weitzman B (2011) The Berber identity movement and the challenge to North African states. Texas University Press, Austin
- Owen J (ed) (2000) Arabic as a minority language. Mouton, Berlin

- Payne R (ed) (1983) *Language in Tunisia*. Publications of the Bourguiba Institute of Modern Languages, Tunis
- Penchoen JG (1968) La Langue Berbère en Tunisie et la Scolarisation des Enfants berbérophones. *Revue Tunisienne des Sciences Sociales*, 1968, pp 173–186
- Romaine S (1989) *Bilingualism*. Blackwell, Oxford
- Saa F (2010) Amazigh studies at the Moroccan university (in Arabic). *Langues et Linguistique* 6: 37–46
- Sadiqi F (1997) The place of Berber in Morocco. *Int J Sociol Lang* 123:7–21
- Sadiqi F (2014) Berber and language politics in the Moroccan educational system. In: Ennaji M (ed) *Multiculturalism and democracy in North Africa*. Routledge, London, pp 81–91
- Salhi Z (2000) Language as identity in colonial and post-colonial Algeria *Offshoot* 3(2):29–44
- Steinkrüger PO, Seifart F (2009) Transfer of derivational morphology without borrowing of stems: Resígaro (Arawakan, Peru) and Chabacano (creole, Philippines). Paper presented at *Morphologies in Contact*, University of Bremen, October 1–3, 2009
- Taifi M (1979) *Le Tamazight au Contact de l'Arabe Dialectal*, Etude Sociolinguistique sur le Parler des Ayt Mgild. Thèse de 3ème cycle, Université de Paris III, Sorbonne nouvelle
- Thomason SG, Kaufman T (1988) *Language contact, creolization, and genetic linguistics*. University of California Press, Berkeley
- Versteegh K (2001) Linguistic contacts between Arabic and other languages. *Arabica* 48(4): 470–508
- Wagner DA (1993) *Literacy, culture, and development: becoming literate in Morocco*. Cambridge University Press, New York
- Weinreich U (1968) *Languages in contact*. The Hague: Mouton. Linguistic Circle of New York, New York
- Youssi A (1992) *Grammaire et lexique de l'arabe marocain moderne*. Wallada, Rabat

Comment la tradition berbérissante a intégré la partie ergative de la syntaxe berbère actuelle dans celle qui est accusative

19

Abdelaziz Allati

Contents

Introduction	403
Des éléments syntaxiques ergatifs intégrés dans la forme prédominante de la syntaxe berbère moderne et actuelle	404
Deux types de verbes et de syntaxe dans les berbères moderne et actuel	415
Conclusion	422
Bibliographie	422

Introduction

Un survol rapide de l'évolution des études syntaxiques et lexico-sémantiques berbères montre qu'elles n'ont pas avancé dans l'analyse de plusieurs éléments centraux dont ne sont traitées que les aspects qui apparaissent du point de vue adopté par la tradition berbérissante. Chaque fois que celle-ci a buté sur ces éléments, elle n'a pas pu sortir de ce cadre et s'interroger sur l'angle de vue adopté et ses fondements. Les berbérissants qui s'y sont aventurés n'en ont ainsi effleuré que la surface se contentant généralement d'interrogations, de présomptions, d'interprétations... (cf. plus bas). La tradition berbérissante a-t-elle fait fausse route? Comment ce point de vue a-t-il remodelé ces éléments et jusqu'à quand va-t-on entretenir, couvrir... le blocage auquel il a conduit? De quel problème s'agit-il au juste et comment replacer ces études sur la bonne voie?

A. Allati (✉)

Université de Tanger-Tétouan, Tanger-Tétouan, Morocco

e-mail: a.allati@uae.ac.ma

Des éléments syntaxiques ergatifs intégrés dans la forme prédominante de la syntaxe berbère moderne et actuelle

La conception pré-bassetienne¹ a décrit le berbère en se basant essentiellement sur les traits structurels de la langue française et, secondairement, sur ceux du sémitique². On appliquait les traits structurels du français aux éléments berbères et la partie qui ne s'y prête pas est ramenée à ceux du sémitique (Allati 2002, 2012, 2015a, b, 2017, 2018, 2020a).

Se basant sur les caractéristiques du verbe en français (ses types/espèces, la processivité fondamentale du verbe. . .) et sur ses types de dérivation en sémitique (cf. Cohen 1924), la conception pré-bassetienne a distingué, d'une part, les verbes d'état des autres (verbes d'action ou de procès) et, de l'autre, la forme simple du verbe de ses formes dérivées (formes passive, factitive, active, réciproque, forme d'habitude, cf. Basset, R. 1895; Destaing 1907, et d'autres). De ce point de vue, la forme de base du verbe en berbère ou sa forme simple ne peut être que le verbe de procès (la partie des verbes dont le premier déterminant est un sujet agent (=V2, cf. groupe C, plus bas): *ešš*, (et var.) « manger », *ney* (et var.) « tuer », *ru* (et var.) « pleurer », pan-berbères, etc., Le reste en serait des formes dérivées, formées à partir de la forme verbale simple comme il en est dans les références française (formes pronominales, passive. . . du verbe) et sémitique (factitif, passif. . .).

La conception bassetienne (cf. Basset 1929, 1952) a opéré une nouvelle redistribution des positions des deux références française et sémitique à partir desquelles était décrit le berbère en mettant cette dernière en première position tout en prétendant avoir totalement rompu avec la seconde (cf. Basset 1929, 1952; Cohen 1993). Dans la réalité, sous le couvert de la description du berbère par ses traits intrinsèques, elle y a étendu – aidée par les correspondances entre ces deux groupes – au maximum les traits structurels sémitiques tout en les maintenant là où ils étaient déjà, et elle a repris, dans les parties où ils ne peuvent être transposés, l'analyse de la conception pré-bassetienne se basant sur la référence française (cf. Allati 2018, chap. I et II). Aussi sa présentation du verbe est-elle une nouvelle reconfiguration des éléments sur lesquels se base la précédente conception (cf. plus bas):

« Le verbe comprend une forme simple et des formes dérivées. La dérivation peut être obtenue (. . .) par l'addition de préfixes à base consonantique, ainsi, à base sifflante, à base dentale, à base nasale, ce qui représente les trois formes vivantes de dérivation, la première à valeur de factitive, la seconde à valeur passive, la troisième soit à valeur réciproque soit à valeur passive, selon les verbes » (Basset 1952: 12).

¹La conception pré-bassetienne est le cadre dans lequel le berbère a été décrit avant les travaux d'André Basset.

²Cela est dû, outre le progrès réalisé dans le domaine sémitique, aux relations historiques/géographiques que ce groupe a avec le berbère dont l'exploration revenait au début aux sémitisants/arabisants.

Cette approche a bloqué les études syntaxiques dont notamment la syntaxe du verbe au temps de Basset comme elle continue de le faire de nos jours:

« De même que nous ignorons la nuance impliquée par la présence du suffixe *t*, il reste encore beaucoup à apprendre sur les raisons d'être et les nuances des formes dérivées (...) Mais surtout pourquoi l'expression du passif par une forme dérivée quand, en base, comme l'a excellemment vu Foucauld, la forme simple par elle-même a généralement les trois valeurs d'actif, passif et réfléchi? Et si parfois certains verbes se refusent aux trois valeurs, c'est la valeur passive ou une valeur voisine, qui est exclusive à la forme simple, l'actif étant alors exprimé, par suite de glissement du factitif à l'actif, par la forme à sifflante. Tel est le cas de *ənz* "être vendu", *ərs* "être posé", etc., "vendre", "poser" s'exprimant par *zzənz* (*ssənz*) et par *ssərs*. Enfin, de façon générale, la forme à sifflante n'a pas que la brutale valeur de factitivité, la forme à dentale de passif et, sous les réserves précédentes, la forme à nasale de réciprocité » (Basset 1952: 13).

Si l'on excepte la forme d'habitude qu'il a intégrée dans le système verbal³, Basset a repris telle quelle l'analyse que la conception pré-bassetienne a faite du verbe et de ses formes sans revoir le type d'analyse qui en a été fait, ni les fondements sur lesquels il se base (cf. plus haut). Il en est de même de plusieurs éléments dont la racine, les préverbes, ce qu'on appelle l'opposition d'état, etc. (cf. Allati 2017, 2018, 2020a).

La conception bassetienne a ainsi repris l'analyse pré-bassetienne des verbes statifs (=V1) dont le premier déterminant est un patient non agent (cf. A et B.1, plus bas). Se basant sur les traits de leurs correspondants dans les traductions en français, sur le type de traitement consacré au statif dans le sémitique⁴ et sur les caractéristiques du verbe de procès (cf. C, plus bas), considéré comme la forme verbale de base en berbère (cf. plus haut), elle les a pris – tout comme la conception précédente – pour des passifs⁵ (cf. plus bas). Ce faisant, elle n'a pu / ne peut voir dans *s*-V1 – de par son sens – que l'actif correspondant, ce qui revient à les ramener à l'opposition de diathèse (cf. plus haut). Et cela tout en accordant une place particulière aux verbes d'état – ou en les traitant à part - en se fondant, outre la distinction de ce type de verbes des autres dans la référence française, sur leurs caractéristiques morphologiques (conservation de la conjugaison suffixale dans

³ « Ce que nous appelons impératif intensif et aoriste intensif était réuni jusqu'ici sous le nom de forme d'habitude, considérée comme une forme dérivée, au même titre que les autres formes dérivées à sifflante, dentale ou nasale, par exemple » (Basset 1952: 14).

⁴ « En fait, à quelque moment de l'histoire de l'akkadien qu'on le saisisse, le statif nous montre un stade dans un processus de verbalisation. Lorsqu'il a pour base un simple nom ou un adjectif non verbal, il est encore une sorte de phrase nominale; mais un thème verbo-nominal de valeur généralement passive » (Cohen 1984: 109).

⁵ On a pris le premier emploi des verbes dont le premier déterminant est soit un patient non agent, soit un agent sujet, pour la valeur passive du verbe simple (cf. B.1, plus bas).

plusieurs variétés: kabyle, touareg, etc.).⁶ On les a séparés/isolés des autres verbes statifs et notamment de ceux dont le premier déterminant est un patient non agent (cf. A.2, plus bas) tout en escamotant leurs traits syntaxiques et lexico-sémantiques communs⁷ dont on ne savait/sait ni comment poser le problème, ni comment le traiter et, encore moins, le résoudre (cf. plus bas).

De plus, si la tradition berbérissante n'a pas réussi à déterminer les fonctions de *s-* déterminant le verbe et le nom (*s-V/(N)*)⁸, c'est qu'elle a abordé cet affixe des points de vue morphologique et extérieur (se basant sur les traits du verbe en français et en sémitique) à partir desquels il a été approché auparavant (cf. plus haut). Basset y a vu un factif qui subit un glissement à l'actif (cf. ci-dessous), alors qu'il s'agit, pour Galand, généralement d'un factif ou causatif⁹. Analyse reprise par Chaker dont le seul "exploit" est d'ordre terminologique. *s-* serait, pour lui, principalement un transitivant/agentivant (cf. Chaker 1995: 73). Il a pris de même *s-*, dans *s-N* (= verbe), pour un verbalisateur, une fonction qui serait secondaire de cet affixe (cf. Ibid.). Ne s'agit-il pas après tout de la même chose que ce qu'il a qualifié de transitivant/agentivant (cf. plus bas)? En somme, il n'a fait que reprendre l'approche de Basset où, à la place du « glissement du factif à l'actif » (Basset 1952: 13, cité plus haut), il a inversé l'ordre de ces deux éléments. Il a privilégié les traits de « l'actif » (transitivité, agent/sujet) reléguant ainsi le « factif » au second plan, et ce en se basant essentiellement sur les traits syntaxiques des correspondants des verbes berbères dans leurs traductions en français. Force est de constater que l'opposition transitif- intransitif¹⁰ ne distingue les verbes simples berbères de leurs formes déterminées par *s-*¹¹ que si on les aborde à travers les traits du verbe en français (cf. plus bas). Le moins qu'on puisse dire est que Chaker n'a fait que remodeler le type de déformation de cet affixe et de la syntaxe du verbe, et que, terminologie mise à part, son analyse n'est pas au fond différente de celles faite dans le cadre de la conception pré-bassetienne (fin du XIXe siècle et début du XXe).

Les études syntaxiques berbères ne sont-elles pas encore à ce stade ou presque? Outre l'énorme fossé entre le verbe statif déterminé par *s-* et l'actif (ou les autres

⁶ « Les descriptions des parlers berbères consacrent généralement un chapitre au groupe des verbes de qualité appelés aussi verbes d'état » (Galand 1955: 245). Dans son article « Continuité et renouvellement du système verbal », Galand a de même subdivisé le point IV.A., l'opposition accompli- inaccompli, en 4.1. verbes de procès et 4.2. verbes d'état (cf. Galand 1977: 293–297).

⁷ « De nombreux verbes pourtant dépourvus de ces caractères morphologiques se comportent dans tous les contextes, par leur valeur et leur emploi, comme des verbes décrits plus haut [les verbes d'état]. Ainsi *arg* « être chaud », *nu* « être cuit », etc. » (Galand 1955: 246).

⁸ Le nom est mis entre parenthèses parce qu'il ne reste que des résidus de *s-N* dans quelques variétés berbères actuelles dont le tachelhit. Les changements syntaxiques que le nom a subis lors de la formation du berbère moderne ont réduit considérablement ses emplois *s-N* (cf. plus bas et Allati 2018, 2020).

⁹ « Les dérivés à sifflante généralement factitifs ou causatifs » (Galand 1988: 234).

¹⁰ « *s-* se combine avec (...) des verbes simples intransitifs de la classe des verbes d'état ou des déponents qu'il rend transitifs » (Chaker 1995: 73).

¹¹ Par exemple, *enz* (et var.) « être vendu » et *zzenz* (<zz- (<-s)- *enz*) « vendre », cf. plus bas).

appellations dont transitivant/agentivant...), on ne sait pas encore ni comment distinguer celui-ci du factitif/causatif et du *s*- “verbalisateur” (cf. ci-dessus), ni quelles relations entretiennent ces trois éléments (cf. plus bas). Car on n’a fait en réalité que reprendre les premiers travaux où sont déterminées non les fonctions de *s*-, mais celles des correspondants de *s*-V/(N) dans leurs traductions en français, et ce sans manquer d’y générer de nouveaux faux problèmes (cf. idem). Comme on n’a pas pu distinguer l’actif du factitif/causatif à partir des traits du verbe en français, on les a, d’un côté, juxtaposés en faisant passer l’un avant l’autre (le factitif avant l’actif (Basset) et inversement (Chaker et autres, cf. ci-dessous) et, de l’autre, on a réduit le premier au second (Galand, cf. note, 9). Et l’on voit que les principales propositions sont restées dans le même cadre où est privilégié l’un ou l’autre pôle,¹² et que, en définitive, tout cela n’y a guère dépassé le niveau terminologique. On ne sait donc de cet affixe que les éléments à partir desquels il est abordé (cf. plus bas). Aussi les études sur la syntaxe du verbe sont-elles restées en grande partie dans le cadre où les a placées la conception pré-bassetienne (cf. plus haut).

Autant d’éléments qui montrent que le problème a été mal posé ou plutôt il est posé de l’extérieur. N’ayant put le voir de l’angle de vue où il s’est placé, Basset s’est limité à s’interroger sur les points qui lui paraissaient incohérents, voire inexplicables au stade où en étaient les recherches (cf. la citation, plus haut). Restés au même angle de vue, la plupart des berbérissants n’y ont vu, tout au plus, que des points qui restent à élucider, voire des pistes à explorer. D’aucuns sont même allés jusqu’à y voir les traits marquants de la syntaxe berbère,¹³ alors qu’il s’agit de formes apparentes du blocage auquel la démarche choisie a conduit les études syntaxiques berbères: le passif exclusif à la forme simple, le double passif, l’actif obtenu par une forme dérivée... (cf. plus bas). Et l’on voit que la déformation de ces éléments est si profonde qu’elle a fourvoyé même ceux qui se font passer actuellement pour des maîtres en la matière. Force est de constater que traiter des problèmes auxquels est confrontée la syntaxe du verbe du point de vue des conceptions pré-bassetienne et bassetienne est une chose, déterminer les traits qui la caractérisent en est une autre (cf. plus bas).

Très peu de berbérissants ont en tout cas revisité les éléments que Basset a signalés (cf. plus haut), qui ont été/sont généralement ignorés, éludés, escamotés... faute essentiellement de recul suffisant pour pouvoir voir les faits comme ils se présentent sans les déformations qu’ils ont subies (cf. plus bas). N’ayant pas réussi à déterminer le problème qui y est posé, Galand (1981, 1987) et Chaker (1995, 1996) n’en ont même pas écorché la surface (cf. Allati 2011, 2013, 2015a, b, 2017, 2018, 2020 et plus bas). Nous ne nous arrêtons ici que sur la proposition la plus récente et la plus “audacieuse”.

¹² Les études d’obédience générativiste se basent en général sur la vision de la tradition berbérissante dont notamment les travaux de Basset dont elles ne diffèrent au fond que par les présentations formelles.

¹³ « Basset mettait là le doigt sur une série de particularités marquantes de la syntaxe du berbère » (Chaker 1995: 63)

Pour Chaker (Ibid.), la valeur passive du verbe simple est plutôt la non-orientation ou la disponibilité du prédicat (cf. plus bas) qui caractérise le prédicat verbal dans les langues ergatives et qui constituerait ainsi les résidus d’une ancienne ergativité de la syntaxe berbère. Plusieurs verbes seraient, selon lui, non orientés par rapport à leurs déterminants (sujet/agent, patient):

« Un verbe comme *kkes* “enlever, ôter”, peut en effet avoir le sens de “être enlevé”, cela sans aucune modification formelle:

- *y-kkes* « il a/est enlevé »

(2/a) - *y-kkes (afrag)* « il a enlevé (la clôture) »,

(2/b) - *y-kkes (wfrag)* « la clôture est enlevée » (Chaker 1995: 64, 1996: 2678).

Pour avoir un même prédicat verbal (*y-kkes*) dans les deux énoncés – l’argument décisif dans cette analyse –, Chaker s’est basé sur la structure de 2/a pour segmenter 2/b où *wfrag* – qui n’est pas attesté isolé – fait partie du syntagme prédicatif verbal composé du verbe et du premier déterminant ayant deux formes: grammaticale (l’indice personnel (IP): *y-*) et lexicale (*w-frag*). Celui-ci est aussi important que le premier dans ce type de prédicat verbal avec lequel il a le même type de rapport (cf. Allati 2020). D’autant plus que, du point de vue diachronique, le premier déterminant lexical (*-wfrag*, ici) était le seul premier déterminant du prédicat verbal dans le stade pré-moderne avant que l’indice personnel n’y fût rajouté lors de la formation du berbère moderne (BM) et notamment du paradigme personnel prédominant dans ce stade et le stade actuel (cf. Allati 2020a, b, 2023). Ce paradigme ne s’est pas pourtant généralisé à tous les verbes et, tout particulièrement, aux verbes d’état dans plusieurs variétés, qui conservent des paradigmes personnels où la marque de la troisième personne est facultative ou absente¹⁴ (cf. plus bas et Allati Allati 2020a, b, 2023). L’antériorité historique du premier déterminant lexical sur celui qui est grammatical est de même corroborée par l’ordre principal des mots dans le phylum afro-asiatique (VSO, cf. Greenberg 1963; Diakonoff 1965, 1988) que partage cette langue¹⁵ notamment avec le sémitique et l’égyptien.

Se basant sur la structure du verbe dans les traductions en français de 2/a et 2/b, Chaker a déformé la structure syntaxique de ce dernier pour lui attribuer un prédicat verbal identique (*y-kkes*) à celui du premier, ce qui prouverait qu’il se caractérise par

¹⁴ La marque de la troisième personne du sing. est facultative dans le paradigme personnel conservé par les verbes d’état au prétérit dans le parler des Ayt-Ziyan. (cf. Galand 1990: 127). Celle de la troisième personne masc. sing. est absente dans les conjugaisons suffixales conservées par les verbes d’état au prétérit dans plusieurs variétés (kabyle (Grande Kabylie), touareg, etc.), absence qui est considérée comme un trait archaïque : « À l’accompli, la troisième personne du singulier conserve une forme archaïque » (Galand 1990: 131).

¹⁵ Pour les berbérissants, VS qui caractérise cet emploi est la forme fondamentale et la plus ancienne de l’ordre des mots en berbère: « D’après l’état du nom (...), on est en droit de penser que, fondamentalement, en berbère, le sujet du verbe est après le verbe à l’état d’annexion et que, en base, ce que l’on considère comme un sujet avant verbe, à l’état libre, n’est qu’une anticipation du sujet » (Basset 1959: 94 ; cf. également Galand 1964).

la non-orientation ou la disponibilité du prédicat qui présenterait des résidus d'ergativité ancienne. Cette déformation est due principalement à deux autres. Tout d'abord, celle de la fonction du premier déterminant lexical du prédicat (*-wfrag*, ici) que la tradition berbérissante a prise pour l'état d'annexion (= EA, cf. Basset 1952: 26) et qu'elle a ainsi réduite à la morphologie nominale (marque du nom¹⁶, cf. Allati 2020a). Ensuite celle que cela a générée et qui prend ce déterminant lexical pour le « sujet du verbe placé après le verbe » (Basset 1952) ou pour une reprise du sujet en se basant sur la structure des traductions de ce type d'énoncés en français¹⁷. Force est de constater que c'est dans ces traductions qu'il y a la reprise du sujet et non dans les énoncés berbères (cf. Allati 2020, et plus bas). Des faux problèmes où se sont enlisées les recherches, si bien qu'on ne sait rien aujourd'hui sur cette fonction du nom, si ce n'est un foisonnement terminologique sans précédent où on n'a fait que reprendre la définition de Basset dans des termes proches ou identiques: « complément explicatif » (Galand 1964)¹⁸, « sujet lexical (explicite) ou reprise lexical du sujet » (Chaker 1995), etc. (cf. Allati, Ibid.). Ayant pris ce qu'on appelle l'état d'annexion pour un trait morphologique du nom dont la fonction serait de reprendre, d'expliquer... le sujet du verbe, Chaker a mis sur le même plan *afrag* et *-wfrag* (EA+ nom) faisant ainsi ramener la structure syntaxique de 2/b à celle de 2/a.

Il s'agit ainsi de deux énoncés qui sont formellement, syntaxiquement et sémantiquement différents [IP-V-*w-frag* (premier déterminant lexical) ~ IP-V-*afrag* (CO)] et qui sont structurellement indépendants, en ce sens qu'ils ne proviennent pas de la non-orientation du prédicat verbal par rapport à ses déterminants, qui aurait permis l'interprétation de ce dernier dans deux sens différents (cf. Allati 2013, 2020b). Il n'y a donc pas de retournement du rôle du premier déterminant et même si c'était le cas, ce n'est pas « sans aucune modification formelle » (Chaker 1995, cité plus haut), ni du verbe, ni du nominal que cela a lieu. Toujours est-il qu'il en est tout autrement dans les prédicats disponibles ou non orientés par rapport à leurs premiers déterminants dans les langues ergatives (sujet/agent, patient)¹⁹.

¹⁶ « Le nom berbère est soumis à trois catégories de variations : l'une intéresse le genre, la seconde le nombre et la troisième ce que nous appelons l'état » (Basset 1952: 23).

¹⁷ « L'énoncé verbal berbère rappelle beaucoup les expressions courantes en français populaire du type : « l'homme, il est venu/il est venu, l'homme ». » (Chaker 1984: 142). Aussi les berbérissants traduisent-ils ordinairement, par exemple, *i-krez u-rgaz* (il-labourer (Pr)-u-homme) par « il a labouré, l'homme, l'homme a labouré » (cf. Galand 1964: 35; Allati 2020b).

¹⁸ Pour éviter le mot reprise, Galand a utilisé plutôt le complément explicatif « qui, possédant par ailleurs tous les caractères du complément déterminatif dont il n'est qu'une variété, ne se confond pas avec le syntagme de reprise » (Galand 1969: 94), mais cela ne change rien à l'analyse.

¹⁹ « Dans ces langues [avar, tongien, chinois, etc.], en gros, l'ensemble prédicat + premier participant suit exactement le même schéma que les énoncés suivants du jeune locuteur francophone; il emploie la même syntaxe pour deux situations préférentiellement opposées: a) /*māžel* bébé ~ b) /*māžel* bonbon; ajoutons-y c) /*māžel* poulet. L'énoncé a) est incompatible avec un nouveau participant agent (...) En b), l'adjonction de l'auteur (...) ne change rien aux rapports existants, mais c) peut se résoudre en deux directions opposées, suivant le sémantisme des expressions: /*māžel* pouletbébé implique poulet victime, mais /*māžel* pouletbonbon vaut pour poulet auteur » (Tchekhoff 1978: 43–44).

Les deux passifs en berbère (Basset 1952, cité plus haut) est un faux problème dû au fait qu'on s'est fondé sur les traductions en français qui font de 2/b la forme passive de 2/a. Ce n'est pas sur les traits syntaxiques de 2/a et 2/b (et des énoncés du même type) que se base l'analyse de Chaker, mais sur ceux les formes active et passive de leurs correspondants respectifs dans les traductions en français, qu'il leur a extrapolés et qu'il a interprétés comme la non-orientation ou la disponibilité du prédicat. Ce faisant, il a fait superposer un nouveau faux problème sur le précédent. La tradition berbérissante a interprété les verbes des énoncés du même type que 2/b comme des passifs que Chaker a pris pour la non-orientation ou la disponibilité du prédicat. On est resté dans les bulles que constituent respectivement la diathèse caractérisant le verbe en français et la non-orientation ou la disponibilité du prédicat dans les langues ergatives et, donc, en dehors des traits syntaxiques du verbe berbère.

2/b est en revanche un emploi statif de *kkes* dont le correspondant le plus proche en français est « il y a eu/il y a état d'enlever la clôture » (cf. plus bas). Et l'on se rend compte non seulement des différences entre 2/a et 2/b qui ne peuvent provenir ni de l'opposition de diathèse, ni du retournement du rôle du premier déterminant du prédicat verbal, qu'on pourrait lui faire correspondre, mais encore, et surtout, du degré de la déformation que ces éléments ont subi.

Qui plus est, le verbe *kkes* « être enlevé, enlever, ôter » sur lequel se base l'analyse de Chaker, n'a que le second sens dans quelques variétés berbères dont notamment le touareg (tahaggart) où il signifie « ôter, enlever... » (cf. Foucauld 1951–1952: 902–906). Ce qu'on a pris pour la non-orientation ou la disponibilité du prédicat verbal par rapport à ses déterminants n'est-il pas en variation dialectale, historique... ? En effet, l'emploi du verbe dans 2/b caractérise un nombre important de verbes (dont les verbes d'état) appartenant généralement au lexique de base et ayant un premier déterminant patient non agent (cf. groupe A, plus bas), mais ne permettant pas l'emploi 2/a. Celui-ci caractérise cependant les verbes de procès dont le premier déterminant est un agent/sujet et qui ne connaissent pas l'emploi 2/b (cf. groupe C, plus bas). La comparaison de ces deux types de verbes avec ceux qui sont, pour Chaker, non orientés par rapport à leurs déterminants (sujet/agent, patient), montre que ceux-ci se situent historiquement entre eux (cf. plus bas et Allati 2011, 2013, 2015b, 2018, 2020b). Evolution qu'illustrent bien des éléments dont notamment les différents stades où en sont plusieurs verbes dont *kkes* dans les variétés berbères actuelles et les traits morphologiques, syntaxiques et lexico-sémantiques qu'ils conservent ou qui y sont fossilisés (cf. plus bas).

Il ne s'agit donc pas de la non-orientation ou de la disponibilité de *kkes*, mais de ses traits lexico-sémantiques et syntaxiques (emploi statif, verbe de procès) caractérisant ses emplois dans deux types de syntaxe différents (Allati, Ibid.). En berbères moderne et actuel, le locuteur avait/a la possibilité d'employer les verbes, qui sont au même stade évolutif que *kkes* dans plusieurs variétés berbères, au statif ou comme verbes de procès (cf. plus bas), ce qui correspond à l'opposition de voix en français et à la non-orientation ou la disponibilité du prédicat dans les langues ergatives. Ce n'est donc pas le retournement du rôle du premier déterminant dû à la non-orientation ou la disponibilité du prédicat verbal, qui distingue 2/b de 2/a, mais

le fait que *kkes* exprime le statif dans le premier cas et le procès dans le second (cf. plus haut). On a pris le statif et le procès caractérisant ces verbes pour l'opposition de diathèse (actif - passif) où Chaker a vu la non-orientation ou la disponibilité du prédicat.

De plus, ce double emploi du verbe dont les traits communs qui le caractérisent (cf. ci-dessus) remonte à la formation du berbère moderne et présente ainsi non les vestiges d'une ergativité ancienne de cette langue, mais les traits d'un nombre important de verbes modernes et actuels (cf. plus bas). Aussi retrouvons-nous une des caractéristiques principales de la conception diachronique de la tradition berbérissante: projeter les éléments modernes et actuels de cette langue sur ses stades antérieurs, ce qui conduit à placer toute son évolution entre ses formes modernes et actuelles qu'on fait évoluer les unes des autres (cf. Allati 2002, 2011, 2013, 2017, 2018). Chaker n'a pas seulement généralisé les traits du type de verbe qui est au stade évolutif statif - procès (cf. B, plus bas) à ceux de tous les autres (cf. ci-dessous), mais encore il en a fait la non-orientation ou la disponibilité du verbe qui serait des résidus d'ergativité berbère ancienne. Notons que le verbe berbère conserve des traits ergatifs, mais il ne s'agit pas de ceux que Chaker a cru y découvrir (cf. plus bas et Allati 2011, 2013, 2017, 2018, 2020b).

Ne s'étant pas détaché de la notion du "passif" et se fondant sur une partie des verbes du berbère actuel, celle dont le premier déterminant est soit un patient non agent, soit agent sujet (cf. B.2, plus bas), Chaker a forcé ce type de verbes pour lui attribuer la non-orientation du prédicat (cf. Chaker 1995: 66-69, 1996) qui permet, comme l'opposition de voix, le retournement du rôle de leurs premiers déterminants (sujet/agent, patient) et expliquerait ainsi la valeur passive du verbe simple en berbère. Il a pensé que celle-ci ne pourrait être que la non-orientation ou la disponibilité du prédicat dont le rôle syntaxique est équivalent:

« M'inspirant des analyses de la construction ergative développées par A. Martinet (1968, 1975), je crois donc que la seule façon d'expliquer ce phénomène (...) est de reconnaître que le syntagme prédicatif berbère est un prédicat d'existence qui entretient avec l'indice de personne (et l'explicitation lexicale de celui-ci) une relation non spécifiée (agent, patient, attributaire) » (Chaker 1995: 69, 1996, mis en italique par nous).

L'objectif étant d'expliquer ces emplois du verbe en berbère et non de les décrire, cela n'a ainsi rien changé dans les études syntaxiques de cette classe/catégorie (ses emplois, ses valeurs aspectuelles, etc.) et, en général, de cette langue. Car Chaker est resté dans le cadre de la conception bassetienne à partir duquel il a interprété les correspondances syntaxiques entre, d'une part, la diathèse caractérisant les traductions en français des énoncés de type 2/a et 2/b et, de l'autre, la non-orientation ou la disponibilité du prédicat dans les langues ergatives. Il n'a fait au fond qu'ajouter un autre prisme à celui à travers lequel ont été/sont scrutés les traits syntaxiques du verbe en berbère (cf. plus haut). Ce faisant, il a fait – ironie du sort – superposer une nouvelle déformation sur une celle qu'il voulait démystifier.

L'écueil principal de cette analyse – qui montre son caractère simpliste et superficiel – est que Chaker a voulu résoudre le faux problème du double passif en

berbère en utilisant les mêmes éléments qui l'ont généré (cf. plus haut) tout en les agrémentant de la non-orientation ou de la disponibilité du prédicat verbal caractérisant le prédicat verbal dans les langues ergatives. Aussi est-on encore loin du niveau d'analyse où l'on peut voir les traits syntaxiques et lexico-sémantiques du verbe berbère sans le remodelage qu'ils ont subi par les conceptions pré-bassetienne et bassetienne (cf. plus haut), et où l'on se rend compte de l'urgence à repenser le point de vue adopté et les fondements sur lesquels il se base.

Par ailleurs, l'approche bassetienne du verbe et notamment de ses traits syntaxiques et lexico-sémantiques (dont le fait de ramener les V1 à V2, cf. plus bas) a conduit inéluctablement au blocage dans l'analyse de l'aspect verbal. Et cela depuis Basset (1952) qui n'a pas réussi à déterminer les valeurs aspectuelles des deux termes de ce qui, d'après lui, constitue l'opposition fondamentale du système verbal berbère: *prétérit* (Pr) - *aoriste*²⁰ (A). La tradition berbérissante s'est enlisée ensuite dans les tentatives de ramener les verbes statifs à ceux de procès et dans les faux problèmes que cela a générés, et ce tout en essayant de sortir de l'impasse où elle s'est trouvée sans le savoir et dont elle ne s'était pas / ne s'est pas encore rendue compte.

Se situant en dehors de l'opposition verbale fondamentale *accompli* - *inaccompli*²¹ que Galand (1977) a substituée à celle dégagée par Basset (1952): *prétérit* - *aoriste* (cf. ci-dessus), l'état/statif qu'expriment les verbes d'état (ou de qualité) au *prétérit* (*accompli*) y est ramené au verbe de procès ou, plus précisément, à sa forme à l'*accompli*.

Tout d'abord, se basant principalement sur l'interprétation dont a fait l'objet l'emploi statif du verbe dans le sémitique dont notamment l'*akkadien*²², Galand a fait de l'état qu'expriment les verbes d'état au *prétérit* un *résultatif*, le résultat d'un procès *accompli* antérieurement:

« L'*accompli* lui-même, dans un système à deux pôles où il se définit par opposition à l'*inaccompli*, ne peut signifier un état que par un biais : comme le résultat d'un procès *accompli* » (Galand 1980: 352, voir également Galand 1955).

²⁰ « Les termes d'*aoriste* et *prétérit* sont des termes arbitraires. Nous n'arrivons pas encore à déterminer à quelle nuance de pensée répond leur opposition (...) Faut-il y voir une opposition déterminé/indéterminé, momentané/duratif, parfait/ imparfait, etc. ou encore, selon les termes généralement adoptés par les arabisants, *accompli/inaccompli*? Peut-être, mais, pour notre part, nous sommes tentés de chercher dans le sens de l'opposition d'un précis et d'un imprécis » (Basset 1952: 13–14; cf. Allati 2018).

²¹ « Lorsque le verbe exprime un état, l'opposition entre *accompli* et *inaccompli* prend un caractère particulier (...) Les *accomplis* des verbes dits de « qualité » ont bien la valeur stative et peuvent même concurrencer victorieusement les constructions nominales » (Galand 1977: 296–297; cf. également Galand 1980, 1990).

²² « En raison du fait que dans chaque verbe il (le statif) s'opposait à une forme nommant le sémantème verbal en tant que procès, il fut interprété comme exprimant l'état résultant, l'aboutissement de ce procès » (Cohen 1989: 61).

L'état/statif y est ainsi secondaire par rapport au procès qui est considéré comme le trait fondamental du verbe (cf. Basset 1929, 1952; Galand 1955, 1977, 1980, 1990).

Ensuite, cette théorie étant infirmée par les faits²³, Galand a proposé une autre où il a replacé la précédente dans le type d'évolution imputé au statif dans le sémitique, qui, étant formé à partir d'un nom/adjectif ou d'un verbe dans l'akkadien et ayant évolué à l'accompli dans notamment le sémitique occidental, est rattaché historiquement au nom qui aurait subi un processus de verbalisation (cf. Cohen 1984, 1989). La verbalisation du nom en berbère serait corroborée, d'une part, par une des formes libyques du nom, *gldt*²⁴ (cf. Chabot 1940), considérée comme une conjugaison "nominale" qui correspondrait à celle du statif formé du nom en sémitique²⁵ (cf. Galand 1990: 132), et, de l'autre, par l'identité des formes des verbes d'état au prétérit et celles du nom et/ou de l'adjectif correspondant(s)²⁶ (cf. Galand 1990: 132, 1980; Allati 2008, 2011, 2015b, 2018, 2020a et plus bas).

L'état/statif qu'expriment ces verbes au prétérit a été placé hors de la sphère du verbe, dans celle du nom (nom = état/statif, cf. Galand 1990). Il proviendrait ainsi de celui caractérisant cette classe/catégorie dont ces verbes seraient issus (cf. Allati 2015a, 2017, 2018, 2020 et plus bas). On a fait de ces verbes des éléments dont le processus de verbalisation n'est pas encore achevé, nécessitant encore de perdre le statif/l'état caractérisant le nom dont ils proviendraient. Les traits archaïques que conservent les verbes d'état ont été ainsi placés, comme dans le sémitique, dans une des phases historiques du nom et seraient ainsi un des traits de son évolution (cf. Allati 2013, 2015b, 2018).

Cela permettrait de rendre compte du statif du verbe d'état en inversant le sens de l'interprétation de la relation historique entre le procès et le statif/état (statif/état > résultatif/procès)²⁷, et tout le reste ne serait qu'un travail d'intégration²⁸:

« Le problème posé en berbère (...) est d'accueillir dans son système verbal un syntagme nominal qui, devenu forme conjuguée, n'en garde pas moins la valeur d'un statif » (Galand 1980: 358, voir également Galand 1990).

²³ « Cette théorie se heurte à la réalité non linguistique, dans laquelle un objet "blanc" n'est pas nécessairement un objet "blanchi". De plus, on ne comprenait pas pourquoi les verbes de qualité présentent une morphologie insolite. Il fallait donc corriger l'analyse... » (Galand 1990: 124).

²⁴ *msnsn gldt* (< Massinissa - roi - lui) « Massinissa, le roi ; le roi Massinissa » (cf. Galand 1990: 132 et plus haut). Cette forme est attestée dans une des inscriptions libyques de Thugga/Dougga dont seule une d'entre elles est datée (138 av. J.-C., cf. Chabot, *Ibid.*).

²⁵ *šarr-â-ku* « roi-je, je suis roi », *šarr-â-tu* « roi- tu (masc.), tu es roi », akkadien (cf. Cohen 1989: 175).

²⁶ *zeggay* = prétérit ; (*a*)*zeggay* = nom/adjectif.

²⁷ « La notion de résultatif garde son rôle central, mais le cheminement va en sens inverse. Au lieu du procès à l'état, on part d'une tournure à valeur stative qui n'était pas un verbe, mais qui, en raison même de l'importance accordée par la langue à la valeur résultative, a pu être interprétée comme un résultatif, donc rapportée à un procès : dès lors, elle devenait un accompli » (Galand 1990: 124).

²⁸ Galand (1980) y a distingué l'intégration des formes de celle des valeurs.

Inutile de chercher les vestiges de cette intégration dans les berbères moderne et actuel, puisqu'on l'a également transférée du sémitique où elle est utilisée pour intégrer le statif (exprimé par le nom et le verbe) dans le système verbal de ce groupe en en faisant des formes en stade de verbalisation ayant la valeur du résultatif ou du constatif (cf. Cohen 1989: 172 et plus haut):

« L'intégration de ce statif dans le système verbal a eu de grandes conséquences pour les langues sémitiques connues à des dates plus récentes que l'akkadien » (Cohen 1989: 61)²⁹.

Pour pouvoir intégrer, dans l'évolution du berbère, cette intégration – transférée du sémitique – de l'état dans le procès, on a étendu, au niveau lexico-sémantique, le type d'évolution morphologique qu'ont subi les verbes d'état qui ont perdu, dans la plupart des variétés, leur conjugaison suffixale au profit du paradigme de conjugaison prédominant dans le stade actuel. Les mêmes pressions analogiques qu'aurait exercées ce paradigme de conjugaison sur celui conservé dans les verbes d'état, auraient été/seraient à l'œuvre au niveau lexico-sémantique faisant évoluer l'état/statif au résultatif³⁰. Au transfert, au berbère, du type d'évolution du statif au procès/résultatif et de son type d'intégration dans le système verbal en sémitique, s'ajoute la transposition du type d'évolution des conjugaisons des verbes d'état au prétérît – et donc de celle de leur morphologie – à leur niveau lexico-sémantique. L'écueil principal de cette approche, la falaise contre laquelle tout cela se brise, est que la valeur des verbes d'état au prétérît est toujours l'état/statif³¹ qui continue, en tant que tel, de bloquer les études aspectuelles où l'on a pris/l'on prend le verbe de procès pour la base du système verbal. Si l'on est resté en dehors de ce statif et de son évolution, c'est qu'on a cherché seulement à en faire un résultatif en se fondant sur le type de traitement qui est consacré à son correspondant en sémitique (cf. ci-dessus) et sur l'évolution de la conjugaison de ces verbes (cf. Allati 2015a, 2018 et plus bas).

L'évolution du statif au résultatif et la question de son intégration dans le système verbal sont des faux problèmes dus au fait qu'on a pris le verbe de procès dans les langues sémitiques anciennes et les formes qu'elles ont connues après l'akkadien, et – sur leur modèle – dans le berbère, pour leurs prédicats verbaux fondamentaux et, corrélativement au niveau aspectuel, pour la base de leurs systèmes verbaux. On a, dans l'un et l'autre cas, étendu le procès à l'état/statif alors qu'ils y sont irréductibles. L'abondante littérature traitant des faux problèmes accumulés de cette façon nous donne l'impression que nous avons progressé dans la connaissance des traits du verbe en berbère alors que cela ne fait que nous en éloigner davantage. Force est de

²⁹ Cette intégration du statif dans le système verbal n'est que l'évolution du statif au verbe de procès qui y est escamotée.

³⁰ « La comparaison des parlers entre eux et, même, parfois, les données fournies par un même parler révèlent une tendance à effacer, dans le verbe de qualité, les traits formels qui le séparent des autres verbes. Cette tendance affecte les deux constituants de la forme verbale : le radical et les indices de personne » (Galand 1955: 354).

³¹ Ce qui rend vraisemblable ou plausible cette intégration en sémitique est que l'évolution du statif au procès y est achevée, ce qui n'est pas encore le cas en berbère (cf. plus bas).

constater que la tradition berbérissante a transféré en bloc, au berbère, et les imprudences méthodologiques caractérisant les études sémitiques et les faux problèmes, voire les blocages qu'elles y ont générés (cf. Allati 2008, 2018).

Etant encore au stade où les a laissées Basset ou presque, les études syntaxiques berbères sont caractérisées par un blocage généralisé. Cela est dû essentiellement au remodelage et à la déformation qu'a subis une partie importante de la syntaxe du verbe que la tradition berbérissante a approchée d'un angle extérieur qui l'a fait intégrer dans le type syntaxique prédominant du berbère actuel³² (cf. plus bas). Ce faisant, elle a dévié les études syntaxiques de leur courant ordinaire pour les mener à l'impasse de laquelle elles ne sont pas encore revenues.

Deux types de verbes et de syntaxe dans les berbères moderne et actuel

Pour démêler cet écheveau et accéder au fond qui permet de voir les traits syntaxiques du verbe en berbère tels qu'ils sont, nous commençons par présenter ses différents types dans les stades moderne et actuel (cf. Allati 2002, 2011, 2013, 2017, 2018, 2020b).

A. Les verbes dont le premier déterminant est un patient non agent:

1. les verbes d'état (ou de qualité): *izwey*, *zwey*. . . « être rouge », *ifsus*, *fses* « être léger ». . ., pan-berbère; ex.
- *y-efsus w-eqzin* (< il- être léger (Pr) - w - chien) « il/le chien est léger, est maigre. . . », berbère nord;
2. les verbes qui expriment le statif et dont le premier déterminant est un patient non agent: *irid* « être lavé », *ilil* « être rincé », *enz*, *nez* « être vendu ». . ., pan-berbère; ex.
- *y-enza u-yenduz* (<il- être vendu (Pr)- u – veau) « il/le veau a été/est vendu », berbère nord.

B. Les verbes dont le premier déterminant est soit (1) un patient non agent, soit (2) un agent sujet: *krez* « être labouré, labourer », *mdel* « être enterré, enterrer », *qqen* « être attaché, attacher », *kkes* « être enlevé, enlever, ôter. . . », . . ., plusieurs variétés berbères; ex.

1. - *y-eqqen w-yyul* (< il- être attaché (Pr)- w- âne) « il/l'âne est attaché », berbère nord;
2. - *y-eqqen ayyul* (<il-attacher (Pr) - âne) « il a attaché l'âne »,
- *y-eqqen w-ergaz ayyul*³³ (< il- attacher (Pr)-w-homme-âne),
« il/l'homme a attaché l'âne », idem.

³²Nous distinguons quatre stades principaux dans l'évolution du berbère : le stade proto-berbère, le stade pré-moderne, le stade moderne et le stade actuel (cf. Allati 2018, 2020b)

³³Les verbes de procès (B.2 et C) ont conservé la détermination lexicale caractérisant leur emploi statif (B.2) ou ils l'ont déjà perdu (C, cf. Allati 2020b et plus bas). Cette conservation est due essentiellement au fait que la place du premier déterminant lexical du prédicat, elle, n'a pas changé.

La tradition berbérissante a vu dans le premier emploi de ces verbes dont le nombre est important une forme passive qu'elle a appelée « la valeur passive du verbe simple » (cf. Destaing 1935; Basset 1952; Prasse 1959, 1974... et plus bas).

- C. Les verbes dont le premier déterminant est un agent sujet: *ney* « tuer », *su*, *sew*... « boire », *eṭṭef* « attraper, tenir, capturer... »... pan-berbère; ex.
- *y-eṭṭef aḍbir* (< il- chasser, capturer, tenir... (Pr) - pigeon) « il a chassé, capturé... un pigeon », berbère nord;
 - *y-eṭṭef w-ergaz aḍbir* (< il- capturer, attraper, tenir... (Pr) - w- homme - pigeon) « il/l'homme a chassé, capturé... le pigeon », idem.

A, B et C se réduisent à deux types de verbes: les V1 (A et B.1), des verbes statifs, d'une part et, de l'autre, les V2, des verbes de procès (B.2 et C) qui sont plus nombreux dans le stade actuel de cette langue et qui y sont considérés comme la forme de base du verbe (cf. plus haut et Allati 2002, 2011, 2013, 2015a, b, 2018, 2020b). La comparaison de ces types de verbes montre que A et C appartiennent respectivement aux domaines de l'état/statif (V1) et du procès (V2), alors que B est au stade intermédiaire entre les deux (cf. ci-dessus). Celui-ci est constitué par la partie des verbes qui a évolué aux verbes de procès (B.2 = V2) tout en conservant leur emploi statif (B.1 = V1). Cela montre le type et le sens de l'évolution que cette langue a connu (l'état/statif > l'état/statif - procès > procès) et les types de réorganisation qu'elle a subis depuis le stade proto-berbère (cf. Allati 2002, 2011, 2013, 2015a, b, 2018, 2020b). Etant trop apparent dans le berbère actuel – vu plusieurs éléments probants qui le montrent – pour passer inaperçu (cf. Galand 1955, 1980, 1990), ce type d'évolution caractérise la famille afro-asiatique dont une des oppositions fondamentales est justement statif - procès³⁴. La comparaison de ses différents groupes montre que l'évolution du premier au second y est à des stades évolutifs différents en allant des langues tchadiques au sémitique où elle est achevée, en passant par les autres groupes dont le berbère (cf. Diakonoff 1965, 1988; Allati 2002, 2011, 2013, 2015a, b, 2018, 2020b). Aussi l'opposition statif (état) - procès constitue-t-elle le pivot de l'organisation et du fonctionnement du verbe en berbères moderne et actuel (cf. plus haut et Allati 2002, 2011, 2013, 2015a, b, 2018, 2020b).

D'après cette évolution, les V1 sont constitués par, d'une part, les verbes d'état (A.1) où l'évolution de l'état/statif au procès n'est pas encore entamée et qui sont ainsi les plus conservateurs³⁵ (cf. Allati 2015b, 2018, 2020b et ci-dessous) et, de l'autre, les verbes statifs (A.2) dont le premier déterminant est un patient non agent et qui, à la différence des précédents, sont à un stade avancé dans leur évolution au verbe de procès. Ayant donné lieu aux groupes B et C (cf. plus haut), cette évolution est lente, voire très lente en raison notamment des éléments structurels internes

³⁴ « There is considerable variety in verbal structures of the different Afrasian languages, but the major part of them is characterized by the distinction, at first, between the categories of "action" and "state" » (Diakonoff 1988: 85).

³⁵ En plus des éléments syntaxiques et lexico-sémantiques, ils conservent des traits morphologiques que les autres verbes statifs (A.2 et B.1) ont perdus.

(cf. Allati 2018 et plus bas). La tradition berbérissante a utilisé notamment les traits morphologiques que conservent les premiers verbes au prétérit (plusieurs formes de conjugaisons suffixales dans plusieurs variétés: kabyle, touareg, etc.) pour les distinguer des seconds (cf. plus haut).

Au niveau morphologique toujours, les verbes d'état au prétérit ont des formes identiques à celles du nom et/ou de l'adjectif correspondant(s) (cf. plus haut). Formés, dans la variété de Djebel Nefousa³⁶ et de celle de Senhaja Srair (cf. Gutova 2021) et Ghomara (El Hannouche 2010; Mourigh 2015), sur une base lexicale qui s'apparente au nom actuel, ils y conservent le type le plus archaïque des prédicats verbaux en berbère exprimant l'état/statif et ayant un déterminant patient non agent et un type de conjugaison constitué par des pronoms personnels indépendants et des suffixes de genre et de nombre (cf. Beguinot 1942; Galand 1980, 1990; Allati 2015a, b, 2018, 2020a, b, 2023 et plus bas). Les traits morphologiques, syntaxiques et lexico-sémantiques qui y caractérisent ces verbes à ce thème (absence de l'opposition verbo-nominale (cf. ci-dessus), vestiges d'une conjugaison ancienne, statif, etc.) sont proches de ceux du stade le plus ancien de cette langue ou le proto-berbère. Le degré d'archaïsme de ce type de verbes au prétérit varie dans le berbère actuel en allant de ces variétés/ parlers à celles/ceux, nombreuses/nombreux, où ils ont adopté le paradigme personnel prédominant en berbères moderne et actuel, en passant tout d'abord par le parler de Ayt Ziyān (cf. plus bas) et, ensuite, par les variétés / parlers où ils ont conservé plusieurs types du paradigme personnel suffixal (kabyle, touareg, etc.). Même s'ils ont perdu plusieurs des traits - notamment morphologiques - des verbes d'état dans plusieurs variétés, les autres V1 se caractérisent par les mêmes traits lexico-sémantiques et syntaxiques (cf. Galand 1955; Allati 2002, 2011, 2013, 2015a, 2017, 2018, 2020b, et plus bas).

Plusieurs éléments dont notamment des traits lexico-sémantiques, syntaxiques et morphologiques des verbes d'état et, en général, des V1 dans les stades moderne et actuel, montrent que le noyau de l'énoncé est, dans le stade proto-berbère, un prédicat d'existence (base lexicale) qui ne fait que poser l'existence d'un état, d'un fait, etc. –ayant la valeur du statif–, et dont le premier déterminant privilégié est un patient non agent. Son déterminant sujet/agent y est indiqué, en cas de besoin, par un affixe casuel: l'ergatif (cf. Allati 2002, 2011, 2013, 2015a, b, 2017, 2018 et plus bas). Le type d'évolution qu'a subi cette langue, la phase où elle en est actuellement... montrent que le niveau lexico-sémantique est, dans ce stade, son noyau central dont dépendent la morphologie et la syntaxe (Ibid.).

La construction syntaxique proto-berbère est donc de type ergatif qui est conservé dans les V1 (verbes statifs) dont les formes correspondantes sont attestées dans le tchadique, le couchitique et l'omotique actuels, et conservées dans l'égyptien³⁷ et les

³⁶ Les prétérits de ces verbes sont, d'après Galand, « d'anciennes formations nominales » (Galand 1990: 126; cf. également Galand 1980: §1.4).

³⁷ « Des langues aussi diverses que l'araméen moderne ou l'égyptien ancien présentent des systèmes verbaux fondés entièrement sur des formations verbo-nominales » (D. Cohen 1989: 107).

langues sémitiques anciennes (akkadien, ougaritique, etc.) et modernes³⁸ (cf. Allati, *Ibid.*). L'importance et l'étendue de ces éléments archaïques sont à la base de l'ergativité de la syntaxe proto-afro-asiatique postulée par Diakonoff (1965, 1988)³⁹ et admise aujourd'hui par la plupart des chercheurs dans ce domaine.

La base de la formation du mot y est une base lexicale monosyllabique dont les voyelles sont radicales (VC, CVC), base que conservent les mono- et biconsonantiques dont le nombre est important dans le stade actuel de cette langue.⁴⁰ Ce type de base de la formation du mot est prédominant dans le tchadique⁴¹ et fréquent dans l'omotique et dans le couchitique, et il conserve des vestiges importants dans le sémitique et l'égyptien (cf. Cohen 1947; Diakonoff 1965, 1988; Allati 2015a, 2017, 2018). Données à partir desquelles sont dégagées les bases de la formation du mot postulées proto-afro-asiatiques (CV et CVC, cf. Diakonoff 1988: 42–56). La formation du mot se base, dans ce stade, sur la composition et la reduplication/redoublement qui sont conservés, à des degrés divers, dans les variétés berbères actuelles. De même, y restent encore d'importants vestiges de l'agglutination qui est le type d'adjonction des affixes aux radicaux dans le stade proto-berbère où les oppositions verbo-nominale, de personne, de rection (sujet/objet) et de diathèse sont absentes (cf. Allati 2015a, 2017, 2018).

Plus avancés sont les traits que conservent les verbes d'état au prétérit dans le parler des Ayt-Ziyan (Petite Kabylie), verbes qui s'apparentent à des noms conjugués et qui se caractérisent par un type de flexions personnelles constitué par les pronoms personnels personnels affixes (cf. Galand 1990). Traits qu'ils partagent avec ce qu'on appelle les auxiliaires verbaux (*tuy/ttuy, aqel, qel, aqa*, etc.)⁴² attestés dans plusieurs variétés actuelles (tarifit, Beni Snous, tamazight, kabyle, etc. cf. Allati 2011, 2013, 2018) et avec plusieurs types de prédicats non verbaux⁴³ (adverbes, interrogatifs, démonstratifs. . .) dans la plupart des variétés berbères. Les nouveaux traits de ces types de prédicats dont leur paradigme personnel de

³⁸ « In Semitic (. . .) there is reason to believe that the original verbal construction was ergative » (Diakonoff 1965: 58).

³⁹ « All of this almost certainly means that the Afrasian languages originally had an ergative construction of sentence, which is still preserved in some Cushitic and Chadic languages » (Diakonoff 1988: 101, cf. également, Allati, *op.cit.*).

⁴⁰ Ils constituent une partie importante du lexique de base (*imi* « bouche », *fus* « main », (*a*)*yil* « bras, avant-bras. . . ») dont notamment un nombre important des V1 (A.2 et B.1, principalement); ex. *nez* « être vendu », *irid* « être lavé », *res/ers* « être posé, se poser » . . .

⁴¹ « In the Tchad languages the biconsonantal roots seem to predominate. In spite of the internal vocalic inflection (which, it must be conceded, is much less developed here than in the other Semitico-Hamitic languages), one can, as a rule, establish a root-vowel both in the nominal and the verbal roots [. . .] Still, it can be supposed that the structure of the Tchad verbal root, biconsonantal with a root-vowel, is presumably an ancient feature » (Diakonoff 1965: 38).

⁴² Ceux-ci expriment également le statif.

⁴³ Ex. *manza-t* ? (<où-lui) « où est-il ? », tachelhit ; *anda-k* (<où-tu) « où es-tu ? », kabyle ; *d-iri-t* (<c'est-mauvais-lui) « il est mauvais », kabyle (cf. Dallet 1982; Galand 1990: 133) ; *aba-t* (<il n'y a pas-lui) « il n'y a plus eu de lui (il n'a plus existé ; il est mort) » (Foucauld 1951–1952: 13).

conjugaison⁴⁴, les rapprochent du prédicat ou du noyau de la phrase dans le stade pré-moderne, forme sous laquelle y est réorganisé le prédicat d'existence proto-berbère (cf. plus haut et Allati 2002, 2011, 2013, 2015a, 2017, 2018, 2020b). Un des traits caractéristiques de cette évolution est la recomposition de ce dernier notamment en verbe et en nom. Outre sa fonction de nominal/substantif, celui-ci y détient un emploi verbal équivalent à celui du verbe avec lequel il partage ainsi la plupart de ses traits morphologiques et syntaxiques (prédication, paradigme personnel de conjugaison, négation, etc., cf. Allati 2002, 2011, 2013, 2015a, 2017, 2018, 2020b). Mais dépendantes du niveau lexico-sémantique (le noyau dur du système proto-berbère) qui n'a pas été recomposé (cf. Ibid. et plus bas), la syntaxe et la morphologie proto-berbères ont survécu à cette réorganisation. De plus, d'après les traits syntaxiques des stades proto-berbère, pré-moderne, moderne et actuel, le prédicat y était en principe non-orienté par rapport à ses déterminants (agent/sujet, patient), mais jusqu'à plus ample informé il n'en reste pas de vestiges ni dans le peu de témoignages écrits – de valeurs inégales – qui remontent jusqu'à l'Antiquité⁴⁵, ni dans le stade actuel (cf. Allati 2018, 2020).

Autant les traits des prédicats (verbes d'état au prétérit et autres) déclinés soit par le paradigme personnel se basant sur les pronoms personnels indépendants et les suffixes de genre et de nombre, soit par celui se fondant sur des pronoms personnels affixes sont proches respectivement du prédicat d'existence dans le stade proto-berbère et de la forme qu'il a prise dans le stade pré-moderne, autant ils sont différents de ceux caractérisant le verbe de procès qui est prédominant dans le stade actuel. D'autant plus que ce type de verbe est le produit de la plus importante – mais qui est également partielle, tardive et inachevée – réorganisation morphologique, syntaxique et lexico-sémantique qu'a connue cette langue lors de la formation du berbère moderne qui a eu lieu en deux stades principaux (cf. plus bas et Allati 2011, 2013, 2018, 2020b). Le berbère pré-moderne s'est recomposé tout d'abord en un premier stade du berbère moderne qui se caractérise notamment par la formation de l'affixe accusatif *s-* qui transforme les prédicats verbal et nominal en verbes de procès (*s-N/V = V2*) et, ce faisant, la syntaxe ergative en syntaxe accusative (Ibid.).

Le fonctionnement de ce type de syntaxe et, tout particulièrement, du prédicat (*V1* ou *N*) transformé en verbe de procès (*s-V1/N = V2*) y a ensuite généré le processus évolutif réorganisant l'état/statif en procès et, donc, le premier stade du berbère moderne en son second stade (cf. Allati 2011, 2013, 2017, 2018, 2020b). L'accusatif *s-* s'est réorganisé dans une partie importante des verbes statifs devenus des verbes de procès (une partie importante des verbes du groupe C) dont une partie est restée en position intermédiaire en conservant leur emploi statif (B.1) à côté du

⁴⁴ *zeggay-iyi* « rouge-moi », « je suis rouge », *zeggay-it* « rouge-elle », « elle est rouge », *zeggay-ihn* « rouge-eux », « ils sont rouges », Ayt-Ziyan (cf. Galand 1990). Il s'agit, pour Galand, d'une conjugaison nominale (Ibid.: 128).

⁴⁵ Aux inscriptions libyques, s'ajoutent les sources arabes médiévales (mots, phrases, toponymes... (cf. El Bekri 1965) et les documents d'origine almohade, ibadite...) et d'autres remontant jusqu'au XVI^e siècle.

nouveau emploi (B.2), formant ainsi le groupe B qui se situe historiquement entre A et C (cf. plus haut). Tout verbe B qui perd le premier emploi (B.1) passe au groupe C. La syntaxe et la morphologie pré-modernes (où sont reconduites celles du proto-berbère, cf. plus haut) ont été ainsi partiellement recomposées respectivement en construction accusative et en morphologie dérivationnelle et flexionnelle interne. Les verbes qui n'étaient pas encore réorganisés étaient transformés en verbe de procès par *s-* comme le sont ceux qui ne le sont pas encore: les V1 et quelques emplois résiduels du nom (cf. plus haut). Un nouveau paradigme personnel de conjugaison caractérisant le verbe de procès (V2) a été parallèlement formé et a été étendu progressivement aux V1. Seuls les verbes d'état conservent encore, dans plusieurs variétés berbères, des types de conjugaisons où sont fossilisés des traits structurels proto-berbères et pré-modernes (cf. Allati 2023).

Le processus évolutif réorganisant l'état/statif en procès dont la syntaxe ergative en celle de type accusatif s'est poursuivi en se basant sur l'évolution individuelle des V1 (A.2) en V2 dans les variétés et les parlers berbères modernes et actuels. Plusieurs verbes statifs transformés, par *s-*, en verbe de procès ont ainsi évolué en verbes de procès (B.2) tout en conservant provisoirement leur emploi statif (B.1, cf. plus haut). Evolution que mettent en évidence plusieurs éléments dont les décalages évolutifs entre les stades où en sont plusieurs V1 dans les différentes variétés berbères actuelles⁴⁶, les emplois fossilisés de ce type de verbe⁴⁷, etc. (cf. Allati 2011, 2013, 2015a, b, 2017, 2018, 2020b et plus haut).

Les V1 conservent donc complètement (A) ou partiellement (B.1) leurs éléments lexico-sémantiques et leurs traits syntaxiques ergatifs. Ils sont des prédicats d'existence qui expriment un état ayant la valeur du statif et dont le premier déterminant est un patient non agent. Vus à partir notamment des traits de leurs correspondants dans leurs traductions en français et de ceux du verbe de procès (V2), ils paraissent comme des passifs et ils ont été pris pour tels (cf. plus haut). Il en est de même du statif en akkadien et dans les autres langues sémitiques (cf. plus haut et note 4). L'histoire de la linguistique montre de même que c'est à cette forme du verbe qu'a été ramenée la construction ergative avant d'en déterminer les caractéristiques et de la reconnaître comme un type de construction syntaxique à part entière (cf. Schuchardt 1896; Dirr 1928; Martinet 1958; Tchekhoff 1978 et autres).

Les V1 (A et B.1: verbes statifs) ont été/sont donc ramenés aux traits de V2 (B.2 et C) [X], au schéma structurel aristotélien sujet-verbe-objet caractérisant la construction accusative, et ce avec toutes les déformations qui s'en ont suivi et les

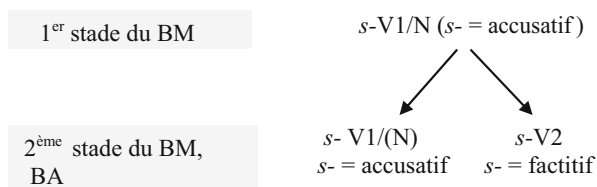
⁴⁶ *nyed/nyed* est au stade A dans le tarifit « être réduit en poudre » (d'où *senyed* « réduire en poudre ») et au stade B dans plusieurs variétés dont le kabyle « être réduit en poudre, réduire en poudre » (cf. Dallet 1982). *yez* (et var.) est au stade A à Ghadamès « être creux, profond » (cf. Lanfry 1973), au stade B « être creusé, creuser » dans la plupart des variétés dont le touareg et le tarifit, et au stade C « creuser » dans quelques variétés dont le parler des Beni Snous, le kabyle (cf. Destaing 1907; Dallet 1982), etc.

⁴⁷ Par exemple, *iri* « vouloir, aimer », pan-berbère, est au stade C dans le berbère actuel, mais son emploi au stade A est fossilisé dans *mezri* « envie, désir » (<*m-z* (<*s-*) -*iri* « être voulu, être aimé »), tarifit (cf. Allati 2002, 2020b).

blocages que cela a générés. Que le berbère n'ait pas, vu le stade où en étaient les recherches en linguistique (cf. ci-dessus), échappé à la confusion du statif avec le passif, cela s'entend, mais rien ne justifie de persévérer dans l'erreur en se cramponnant à une tradition/conception descriptive désuète qui a déformé profondément ses niveaux syntaxique et lexico-sémantique. Une telle déformation a naturellement conduit à d'autres qui ont enchevêtré inextricablement les faits. Si on n'est pas allé au-delà du problème apparent (passif = la forme simple du verbe; l'actif = forme dérivée) signalé par Basset (1952), c'est qu'on ne pouvait pas / ne peut pas en dire plus de l'angle de vue où la tradition berbérissante s'est placée.

Le type d'adjonction de *-s* au verbe (*s-V*) et, surtout, les traits des verbes en français et en sémitique sont à l'origine de la transposition de la relation entre la forme simple du verbe et celle qui est marquée au niveau morphologique (cf. plus haut). Cela a réduit les fonctions de ce préfixe aux traits morphologiques de cette classe/catégorie en en faisant une simple dérivation alors que, outre l'expression de la factitivité, il transforme *V1/(N)* en *V2* et la construction ergative en construction accusative (cf. plus bas), et opère ainsi aux niveaux lexico-sémantique et syntaxique. Ses fonctions ont été/sont ainsi morphologisées et simplifiées au point qu'elles sont devenues totalement méconnaissables. Cela a ainsi laissé dans l'ombre une partie importante de la syntaxe berbère qui ne cadre ni avec les références française et sémitique, ni avec les traits prédominants de la syntaxe berbère moderne et actuelle. Etant générée par la conception qui a pris le verbe statif (*V1*) pour un passif (cf. ci-dessus), la réduction de l'accusatif *-s* à la morphologie verbale a permis de ramener et d'intégrer cette partie de la syntaxe berbère dans celle qui est accusative et prédominante dans les stades moderne et actuel tout en y faisant ainsi aplatis les traits syntaxiques et lexico-sémantiques du verbe et les types de constructions syntaxiques qui le caractérisent.

Les faits apparaissent en revanche tout autrement du point de vue interne. Conjointement à l'évolution d'une partie de *s-V1/N* à *V2* lors de la formation du deuxième stade du berbère moderne (cf. plus haut), l'affixe *s-* a acquis une nouvelle fonction. Il a conservé sa fonction d'accusatif quand il détermine *V1/(N)* qu'il transforme en verbe de procès faisant ainsi basculer la construction ergative dans celle qui est de type accusatif, et il exprime le factitif/causatif en déterminant le verbe de procès (*V2*) qui en a évolué:



Aussi la détermination de ses fonctions est-elle tributaire de celle des traits lexico-sémantiques et syntaxiques (statif ou procès) du prédicat qu'il détermine, sans laquelle on reste au stade de tâtonnements qu'illustrent bien les différentes tentatives que nous avons vues (cf. plus haut). Et l'on commence ainsi à voir la forêt que nous a cachée la vision de la tradition berbérissante.

Conclusion

Ayant repris, sans réexamen préalable, l'analyse pré-bassetienne de la syntaxe du verbe qui a fait intégrer la partie ergative de la syntaxe berbère dans celle qui est accusative en y appliquant notamment les traits structurels du verbe en français et en sémitique, la tradition berbérissante s'est égarée et a conduit les recherches syntaxiques au blocage qu'elle n'a fait qu'entretenir savamment pendant presque un siècle tout en croyant être sur la bonne voie. Ce qu'on découvre derrière le remodelage pré-bassetien et bassetien de la syntaxe du verbe, montre l'ampleur des déformations que celle-ci a subies et les dégâts que cela y a occasionnés. On a cru et on a fait croire que ce sont les traits syntaxiques du verbe en berbère qui sont insaisissables alors qu'ils ont été déformés par ceux du verbe en français et en sémitique (actif, passif. . .) à travers lesquels on les a analysés. On s'en est pris aux éléments décrits, et pourtant c'est le type de regard qu'on a porté sur eux qui est le responsable. Le berbère actuel se caractérise par deux types de prédicats verbaux (verbe statif et verbe de procès) qui sont à la base du système aspectuel et qui correspondent à deux types de syntaxe, ergative et accusative, dans le stade statif (état) - procès où en est son évolution (cf. Allati 2018). Et l'on voit combien nos connaissances sur la syntaxe de cette langue sont, du point de vue de la tradition berbérissante, rudimentaires!

Bibliographie

- Allati A (2002) Diachronie tamazight ou berbère. Publications de l'Université Abdelmalek Essaâdi, Tanger
- Allati A (2008) Proto-berbère et proto-afro-asiatique: l'aspect. In: Takács G (ed) *Semito-Hamitic (Afroasiatic) Festschrift for A.B. Dolgopolsky and H. Jungtraithmayr*. Dietrich Reimer, Berlin, pp 19–26
- Allati A (2011) De l'ergativité dans le berbère moderne. In: Di Tolla AM (ed) *Studi Africanistici, Quaderna di Studi berberi e Libico-berberi I*. Biblioteca Nazionale, Napoli, pp 13–25
- Allati A (2012) From proto-Berber to proto-Afroasiatic. In: Zuckermann G (ed) *Burning Issues in Afro-Asiatic Linguistics*. Cambridge Scholars Publishing, Cambridge, pp 62–74
- Allati A (2013) La réorganisation de l'ergativité proto-berbère: de l'état à l'état/procès. In: Mengozzi A, Tosco M (eds) *Sounds and Words through the Ages: Afroasiatic Studies from Turin*. Edizioni dell'Orsa, Alessandria, pp 177–190
- Allati A (2015a) La place de la composition dans la morphologie berbère. In: *Actes du colloque international: La mesure du sens et le sens de la mesure*. Publications de la faculté des lettres et des sciences humaines Fes-Saïs, Fès, pp 25–45. <https://www.academia.edu/15391685>
- Allati A (2015b) Les verbes d'état berbères: des fossiles linguistiques revisités. In: Di Tolla AM (ed) *La lingua nella vita e la vita delle lingua, Itinerari e percorsi degli studi berberi*, Studi Africanistici, Quaderni di Studi Berberi e Libico-berberi, vol 4, pp 87–116
- Allati A (2017) La dérivation dans la morphologie berbère. In: Allati A (ed) *Auréoles berbères, Mélanges offerts à M. Peyron*. Köppe Verlag, Köln, pp 163–194
- Allati A (2018) L'aspect en berbère, plus de dix mille ans d'histoire d'une langue. L'Harmattan, Paris
- Allati A (2020a) Sur les genèses des marques initiales du nom en berbère: genre, nombre et état. *Études et Documents Berbères* 43:167–185

- Allati A (2020b) Le nom, premier déterminant lexical du prédicat en berbère. *Études et Documents Berbères* 44:135–162
- Allati A (2023) Les paradigmes personnels du berbère: formation et évolution, à paraître
- Basset R (1895) Étude sur la Zenatia de l'Ouarsnis, et du Maghreb central. E. Leroux, Paris
- Basset A (1929/2004) La Langue berbère. Morphologie. Le verbe: étude de thèmes. Réédité par L'Harmattan, Paris/Leroux [2004]
- Basset A (1952) La langue berbère. Oxford University Press, London
- Basset A (1959) Articles de dialectologie berbère. Klincksieck, Paris
- Beguino F (1942) Il berbero nefusi di Fossato. Istituto per l'Oriente, Roma
- Chabot J-B (1940) Recueil des inscriptions libyques. Imprimerie nationale, Paris
- Chaker S (1984) Articles en linguistique berbère. Editions du CNRS, Paris
- Chaker S (1995) Linguistique berbère, Étude de syntaxe et de diachronie. Peeters, Paris/Louvain
- Chaker S (1996) Ergatif (ergativité, non-orientation du verbe). *Encyclopédie berbère*, vol XVII. Edisud, Aix-en-Provence. (reprise du chapitre 6 de Chaker (1995) Linguistique berbère, Étude de syntaxe et de diachronie. Peeters, Paris/Louvain, pp 63–82).
- Cohen M (1924) Les langues chamito-sémitiques. In: Meillet A, Cohen M (eds) *Les langues du monde*. Champion, Paris, pp 81–151
- Cohen M (1947) Essai comparatif sur le vocabulaire et la phonétique chamito-sémitique. Champion, Paris
- Cohen D (1984) La phrase nominale et l'évolution du système verbal en sémitique. Étude de syntaxe historique. Peeters, Paris/Leuven
- Cohen D (1989) L'aspect verbal. PUF, Paris
- Cohen, D. (1993). Racines. In J. Drouin and A. Roth (Éds.), *À la croisée des études libyco-berbères, Mélanges offerts à P. Galand et L. Galand* (pp. 162-175). Paris: Gueutner.
- Dallet J-M (1982) Dictionnaire kabyle-français. SELAF, Paris
- Destaing E (1907) Étude sur le dialecte berbère des Beni Snous. Leroux, Paris
- Destaing E (1935) Note sur le verbe passif. *Comptes rendus du Groupe Linguistique d'Études Chamito-Sémitiques*, T.II. Gueutner, Paris, pp 31–32
- Diakonoff IM (1965) Semito-hamitic languages, an essay in classification. Nauka, Moscou
- Diakonoff IM (1988) Afrasian languages. Nauka, Moscou
- Dirr A (1928) Einführung in das Studium der kaukasischen Sprachen. Asia major, Leipzig
- El Bekri A-O (1965) Description de l'Afrique septentrionale (trad. De Slane MG). Maisonneuve, Paris. [1^{ère} édition, 1911–1913].
- El Hannouche J (2010) Arabic influence in Ghomara berber. Mémoire de Master, Leiden University
- Foucauld Ch. de (1951–1952) Dictionnaire touareg-français (dialecte Ahaggar), vol 4. Imp. Nationale, Paris
- Galand L (1955) État et procès. *Hesperis* 42:245–251
- Galand L (1964) L'énoncé verbal en berbère. Étude de fonctions. In, *Cahiers de Ferdinand de Saussure* 21:33–100
- Galand L (1969) Types d'expressions nominales en berbère. *Cahiers Ferdinand de Saussure* 25:83–100
- Galand L (1977) Continuité et renouvellement d'un système verbal: le cas du berbère. *Bulletin de la Société linguistique de Paris* 72(1):275–303
- Galand L (1980) Une intégration laborieuse: les "verbes de qualité" du berbère. *Bulletin de la Société de Linguistique de Paris* LXXV(1):347–362
- Galand L (1981) Le système verbal berbère et Problèmes de l'orientation du procès en berbère. In: *Actants, voix et aspects verbaux*. Presses de l'université d'Angers, Angers, pp 121–128
- Galand L (1987) Redistribution des rôles dans l'énoncé verbal en berbère. *Actances* 3:132–158
- Galand L (1988) Le berbère. In: Perrot J, Cohen D (eds) *Les langues dans le monde ancien et moderne*, III, les langues chamito-sémitiques. éditions CNRS, Paris, pp 207–242
- Galand L (1990) Du nom au verbe d'état. Le témoignage du berbère. In: Mukarovskij HG (ed) *Proceedings of the 5th International Hamito-Semitic Congress. 1987. vol I. Hamito-Semitic, Berber, Chadic*. Wien, pp 123–137

- Greenberg JH (1963) Some universals of grammar with particular reference to the order of meaningful elements. In: Greenberg JH (ed) *Universals of language*. The MIT Press, Cambridge, MA, pp 73–113
- Gutova E (2021) *Senhaja Berber Varieties: Phonology, Morphology, and Morphosyntax (a polylectal Grammar of Senhaja Berber varieties, Northwestern Morocco)*. Thèse de Doctorat. Univ. Paris III
- Lanfry J (1973) *Ghadamès -II- Glossaire (parler des Ayt Waziten)*. Le Fichier Périodique, Alger
- Martinet A (1958) La construction ergative et les structures élémentaires de l'énoncé. *Journal de Psychologienormale et pathologique* 3:377–392
- Mourigh K (2015) *A grammar of Ghomara Berber (North-West Morocco)*. Thèse de Doctorat. Leiden University
- Prasse KG (1959) Analyse sémantique des verbes dérivés par préfixes en touareg. *Acta Orientalia*, XXIV, pp 147–160
- Prasse KG (1974) *Manuel de grammaire touarègue (tahaggart)*. Vol. VI–VII. Verbe. Akademisk forlag, Copenhagen
- Schuchardt H (1896) Über den passiven Charakter des Transitives in den kaukasischen Sprachen. In: *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien*. Tempsky, Wien
- Tchekhoff C (1978) *Aux fondements de la syntaxe: l'ergatif*. PUF, Paris

Conditionals in Berber

20

Ahmed Ech-Charfi

Contents

Conditionals and Conditional Typology	427
On Berber and Its Verbal System	429
Marking and Clause Order in Berber Conditionals	432
Degrees of Hypotheticality	434
Some Related Adverbial Constructions	439
Conclusion	440
References	440

Abstract

This chapter provides an analysis of conditional constructions in Berber, with a focus on varieties of the language spoken in central Morocco. Although there are relatively recent and more fine-grained typologies (e.g., Dancyger 1998; Dancyger and Sweetzer 2005), we have relied on Comrie's (1986) and Athanassiadou and Dirven's (1995) classifications to consider the extent to which they can account for Berber conditionals. This is mainly because these classifications are based on both formal and semantic aspects while the more recent ones tend to put more emphasis on semantics. In this respect, the chapter considers four parameters, namely, clause order, marking of conditionality, degrees of hypotheticality, and time reference. Since Berber belongs to the Afro-Asiatic language family, and therefore differs in a number of respects from the better-studied European languages, a comparison of Berber conditionals with their counterparts in some of these languages (mainly English) will contribute to some of the different ways in which degrees of hypotheticality are expressed. More specifically, the chapter considers the way the interaction between aspect and time reference as well as the use of different conjunctions

A. Ech-Charfi (✉)
Mohammed V University, Rabat, Morocco
e-mail: a.echcharfi@um5r.ac.ma

affect sentence meaning. The chapter reaches the following conclusions regarding marking, clause order, and degrees of hypotheticality: (1) with some conjunctions, only the protasis is marked, while with others, both clauses are marked. (2) When only the protasis is marked, the order of clauses is relatively free, but when the apodosis is also marked, the order is fixed, with the protasis coming first. (3) When neither clause is marked, the order is also fixed, as before. (4) Constructions in which both clauses are marked express a high degree of hypotheticality than those in which only the protasis is marked. (5) Counterfactuality is not formally distinct from high hypotheticality. (6) Aspect may be exploited to express various degrees of probability, but conjunctions play a bigger role. (7) Within marginal types, wishes are more closely related to conditionals than concessives. However, given the variation, both between and within Berber languages, it is likely that these conclusions may not hold for all varieties.

In his introduction to the special issue of *Studies in African Linguistics*, dedicated to conditionals in African languages, Nicolle (2017: 1) notes that “In contrast to the extensive literature on conditionals in English and other major European languages [...], far less work has been done on conditional constructions in other languages.” Similarly, during more than a century of Berber studies, conditionals in the varieties of this language have received little attention. Neither early grammar writings such as Basset (1929), Destaing (1920), Laoust (1924), nor recent ones such as Bouhkris et al. (2008), Sadiqi and Ennaji (2004), Quitout (1997), to name but a few, make any reference to conditionals. Only Chafik (1991) devotes a few pages to the topic, though his discussion is couched in a terminology developed by Arabic grammarians centuries before the birth of the linguistic science. This in itself constitutes a good reason for a linguistic investigation of conditional constructions in Berber.

Conditionals have been studied extensively by logicians before linguists showed interest in their morpho-syntactic, semantic, or pragmatic aspects. Yet, as late as the 1990s, Athanassiadou and Dirven (1995: 609) complained that “there is as yet no extensive description or analysis of a possible typology of conditionals”. Such a typology seems indispensable if comparison between conditional constructions, both within a language and cross-linguistically, is to be justifiable. Constructions often referred to as conditionals are heterogeneous and can vary widely in their syntactic, semantic, and pragmatic aspects.

Section “[Conditionals and Conditional Typology](#)” below is devoted to the elucidation of these issues. Section “[On Berber and Its Verbal System](#)” introduces the Berber language and its verbal system. Section “[Marking and Clause Order in Berber Conditionals](#)” discusses clause order and clause marking in conditional constructions. Section “[Degrees of Hypotheticality](#)” is devoted to degrees of hypotheticality. Finally, Section “[Some Related Adverbial Constructions](#)” briefly considers concessives and wishes as two marginal types of conditionals.

Conditionals and Conditional Typology

Conditional constructions are usually constituted of two clauses: a subordinate clause (known in English as *if*-clause) stating a condition or a hypothetical situation and a main clause stating a proposition whose truth depends on the truth of the proposition expressed in the subordinate clause. Following the general practice in logic, the subordinate clause will be called the ‘protasis’ and will be symbolized by *P*, while the main clause will be called the ‘apodosis’ and will be represented by *Q*. This notation, however, should not be interpreted as a commitment to the logical approach, which treats conditionals as essentially material implication, a view adopted by Comrie (1986), for example. Unlike material implication concerned only with the truth value of the conditional proposition as determined by the truth table of propositional logic, conditionals in natural language can have a variety of meanings and functions. It is for this reason that different typologies have been proposed so that adequate cross-linguistic comparisons and generalizations can be made.

One of the early attempts to identify the parameters of conditionality is Comrie (1986). As has just been pointed out, Comrie (1986) considers that material implication forms the semantic core of conditionals in natural language and that other meanings usually associated with these constructions are best treated as conversational implicatures, that is, falling within the realm of pragmatics. Although this claim is disputable, Comrie identifies four parameters that could facilitate cross-linguistic comparison. These are (a) clause order, (b) marking of conditionality, (c) degrees of hypotheticality, and (d) time reference. As for (a), he notes that the preferred order is that in which the protasis precedes the apodosis. Comrie also notes in relation to (b) that the most common tendency for languages is to mark the protasis with conjunctions such as the English “*if*” and to leave the apodosis unmarked. In relation to (c), he points out that hypotheticality forms a continuum and that languages vary as to the number of levels or degrees that are grammaticalized. Finally, Comrie notes in connection with (d) that the most common strategy seems to be the use of future time reference with low hypotheticality conditionals and past time reference with high probability conditionals. Comrie (1986), however, does not advance any categorization of conditionals on the basis of these parameters.

In comparison, Athanasiadou and Dirven (1997) argue that conditionality, hypotheticality, and counterfactuality are different semantic categories, though they are often intermingled. For example, although conditionality is typically associated with conditional constructions, it can also be expressed by other constructions. A case in point is the construction “*whenever . . .*” or “*as long as . . .*”, which are generally not considered conditional in English grammar. On the basis of this distinction and the interaction between the three “conceptual domains”, Athanasiadou and Dirven (1997) identify three types of conditionals: (a) course of events conditionals, (b) hypothetical conditionals, and (c) pragmatic conditionals. These are discussed briefly below.

Of the three types, hypothetical conditionals (b) stand as the most typical of this category of constructions, and for that, they are the type investigated most in logic as well as in the linguistics literature. The following example from Athanasiadou and Dirven (1997: 65) illustrates this category:

1. *If there is no water in the radiator, your engine will overheat immediately.*

In this example, the truth or falsity of P and Q simply do not arise because both of them describe hypothetical rather than real situations. That is why material implication tries to take care of the four possible situations in which both P and Q are either true or false or one of them is true while the other is false; the truth value of the proposition as a whole will depend on the truth value of its constituents. More specifically, the conditional proposition would always be true except when P is false and Q is true. The causal relationship between P and Q makes this judgment in accordance with intuition. In (1), for instance, lack of water in the radiator is understood to be the direct cause of the engine overheating; therefore, if there is water in the radiator, the engine would not overheat. This inference, however, is not valid since overheating can be caused by factors other than lack of water. This is a good reason why the conditional cannot always be equated with material implication.

In comparison, the relation between P and Q in (a) is rather weak. The following sentence from Athanasiadou and Dirven (1995: 619) is an illustration of course of events conditionals:

2. *If the tonsils are removed, the adenoids are sometimes cut out, too.*

In this example, the removal of tonsils is not claimed to be the cause of cutting out the adenoids; rather, the relation between them is “that of a natural occurrence or a natural course of events, most closely related to a “whenever” relationship (Athanasiadou and Dirven 1997: 63). This is especially clear from the use of the adverb “*sometimes*”, which indicates that the second event is not entailed by the first. Therefore, the proposition expressed by the *if*-clause could be false, while that in the main clause is true, and vice versa. From a linguistic perspective, the use of the present simple in the main clause is an indication that this type of conditionals is different from hypotheticals, which tend to avoid this verb tense.

In the third type termed “pragmatic conditionals” by Athanasiadou and Dirven (1997), the relationship between the antecedent and the consequent is of a different kind. This is clear from the following example taken from the same source as before:

3. *If you are thirsty, there is beer in the fridge.*

Here, Q is not intended to be a statement in the first place, but rather to invite the hearer by suggesting a solution to his possible problem hypothesized in P. That is why Q often occurs in the imperative mood in pragmatic conditionals and that is what makes this type a special category of conditionals in comparison with course of

events and hypothetical conditionals. Not much will be said about this type in Berber.

Athanasiadou and Dirven (1997) claim that conditionals form a continuum depending on the speaker's commitment to the reality of events and their likelihood of occurrence. But they argue that the possibilities in this connection can be classified into three categories: (a) the speaker commits himself to the real or regular occurrence of P and Q, as in course of events conditionals, (b) the speaker distances himself from any commitment to the reality of P but considers the possibility of Q, as in hypotheticals, and (c) the speaker views both P and Q as unreal, as in counterfactuals. The English language provides a variety of resources to express different degrees of commitment to reality or likelihood of occurrence, thus resulting in a cline ranging from (a) to (c).

In a more or less similar vein, Thompson et al. (2007) distinguish between real and unreal conditionals. Real conditionals include cases in which the event referred to is happening in the present or happens regularly or happened in the past. Accordingly, real conditionals can be further classified into three sub-types depending on time reference. By contrast, unreal conditionals are used when the situations referred to, whether past, present, or future, are only imagined. Thompson et al. (2007) classify unreal conditionals into imaginative and predictive: the first includes those that refer to hypothetical or counterfactual situations, while the second includes those that predict future events.

It should be noted that, although these typologies are intended for cross-linguistic comparison, they are based on data drawn mainly from the major European languages. Obviously, a more adequate typology can only be devised after a wider survey of world languages. Despite that, the proposed categories can serve as a framework to approach conditionals in other languages such as Berber. But before we embark on the description of Berber conditionals, we will first provide a brief presentation of the language and its verbal system because much of what will be said later will hinge on this discussion.

On Berber and Its Verbal System

The title of this paper is rather too broad because Berber is not a single language but a collection of languages scattered across North Africa and some Saharan and sub-Saharan countries. Basset (1952: 1) describes it aptly:

C'est éminemment un ensemble de langues locales, moyen d'expression de groupes sociaux très limités, de quelques milliers, voir même de quelque centaines d'individus, avant tout pour leur usage interne, accessoirement pour leurs relation de groupe à groupe en voisinage immédiat, exceptionnellement à grande distance [It is rather a set of local languages that serve as a means of expression for very limited social groups of a few thousand, or even of a few hundred, individuals, basically for their internal use, incidentally for their immediate inter-group relations, and, only exceptionally, for distant relations]

As a result of this long term sociolinguistic situation, each variety has developed its own lexical, phonological, and grammatical systems, though diachronic relations between the different systems can hardly be denied. But from a synchronic perspective, it would be completely inaccurate to study a grammatical aspect such as conditional constructions as if they all fit within a single Berber system. For this reason, the focus in this paper will be the variety spoken natively by the author, namely, the Tamazight of the Zemmour plateau and its neighbouring dialects of the Middle Atlas in Morocco.

As is the case with most Berber varieties, Tamazight distinguishes between three stems of the verb, each with a different aspectual value: (a) an unmarked stem, (b) an intensive stem, and (c) a perfect stem (cf. Penchoen 1973). These are illustrated by the examples in the following table. It should be noted, however, that these are not exhaustive of the verb forms in the variety described here nor in Berber in general; only the forms that will occur in the examples below are listed. (For more on this issue, see the references in the bibliography):

4.	Unmarked	Perfect	Intensive	Gloss
	<i>krəz</i>	<i>krəz</i>	<i>kərrəz</i>	plough
	<i>səɣ</i>	<i>sya</i>	<i>ssəɣ</i>	buy
	<i>amz</i> ^ε	<i>umz</i> ^ε	<i>ttamz</i> ^ε	catch

As can be noticed, the distinction between the unmarked and the perfect stems may sometimes not be morphologically expressed, as is the case with '*krəz*'. The examples also indicate that different processes are used to derive the surface forms. But these processes, to which grammar books devote lengthy chapters, should not concern us here since they have no bearing on the aspectual meanings; nor should we give any importance to the changes that verb forms undergo in conjugation. Similarly, the distinction between basic and derived verbs (e.g., simple vs. causative vs. reciprocal), which are also given due attention by grammarians of the language, is also not relevant to us, as both classes can express the same aspectual categories.

What is worthy of attention, however, is the aspectual categories themselves. In this respect, the distinction between the perfect and the imperfect stands out as the major opposition in the language (see Basset 1952; Chaker 1989; to cite but a few). The perfect stem is used to indicate that the action is completed, at least from the speaker's point of view, as illustrated by the following examples:

5. a. i-sya urjaz θamazirθ
 3ms-buy.PER man land
 (The man bought the land)
- b. ittuhəl i-sya urjaz θamazirθ
 already 3ms-buy.PER man land
 (The man had bought the land)
- c. asgg°as d-iddan ðað j-ili i-sya θamazirθ
 next year FUT 3ms-be.UN 3ms-buy.PER land
 (By next year, he will have bought the land)

(The following abbreviations will be used: PER = perfect; UN = unmarked; INT = Intensive; FUT = future particle; PART = particle; CON = conjunction; PAS = passive; m = masculine; s = singular; p = plural)

(The morphological gloss used throughout the paper is a simplified one; only when the internal structure of a word is deemed relevant to the discussion will we include it in the gloss)

In (5a), the perfect form indicates that the action of buying is complete and, by implication, corresponds to the past tense in English. In (5b) the form expresses the same idea except that the use of the adverb *‘ittuhal’* (already) indicates that the action had been complete at a certain point in the past, a meaning expressed by the past perfect in English. Similarly, the use of the same perfect form with the verb *‘ili’* (to be) and the future marker *‘dað’* corresponds to the future perfect in English.

An exception to the above remarks is constituted by state or “qualitative” verbs. These generally denote states or qualities that last for a long period of time, as in the following example:

6. i-məqqur / i-zur / i-ɣʷil / urba
 3ms-be big.PER/ 3ms-be fat. PER / 3ms-be good.PER boy
 (The boy is big / fat / good)

Although these verbs are in the perfect form, they do not refer to past actions. Actually, they do not denote actions, but correspond to adjectival predicates in English and similar languages, as is obvious from the English translation.

On its part, the imperfect aspect is expressed by the unmarked and the intensive stems, together with other particles. The most prominent of these particles are the future *‘dað’*, the “conditional, prospective, subjunctive, etc.” *‘að’* (Penchoen 1973: 29); and the extensive *‘la’*. The first two occur with the unmarked stem, while the third one occurs with the intensive stem. These are illustrated by the following examples:

7. a. ðað i-mɣər ijran nnə-s
 FUT 3ms-harvest.UN field.PL of-his
 (He will harvest his fields)
 b. i-ra að i-mɣər ijran nnə-s
 3ms-want PART 3ms-harvest.UN field.PL of-his
 (He wants to harvest his fields)
 c. i-lla la i-mɣɣɣər
 3ms-be.PER PART 3ms-harvest.INT
 (He was harvesting)

In the absence of other indicators, the future reading of (7a) is the most natural reading. The particle *‘dað’* seems to be restricted in dialectal distribution since most other varieties, including that of the neighbouring Ayt Ndhir, described by Penchoen (1973), use *‘að’* to refer to future actions. In the Zemmour variety, *‘að’* is limited to

subordinate clauses, as in (7b), which is ‘*to*’ as a marker of the infinitive in English. As to (7c), the extensive particle, together with the intensive stem, expresses progressiveness, continuity, iterativeness, or generality. The verb ‘*ili*’ (to be) in the perfect form indicates that the action was in the past; without it, the sentence would be referring to a present action.

In brief, Berber languages in general, and the variety under description in particular, have a verbal system that is basically aspectual. Other particles, together with contextual clues, including adverbials, can be used to express a specific time reference. These resources are exploited in conditional constructions to which we turn in the remainder of this paper.

Marking and Clause Order in Berber Conditionals

As is the case in most languages of the world, the protasis and the apodosis tend to follow a certain order in Berber conditionals. In addition, Berber conditionals exhibit the same clause marking noticed in conditionals cross-linguistically.

Concerning clause order, it seems to be more usual for the protasis to precede the apodosis. This order is exemplified by (8a) below, though the opposite order is also attested, as illustrated by (8b):

8. a. [*xm* i-wwəθ unz^oar^o]_P [ðað i-qqim g uxam]_Q
 if 3ms-hit.PER rain, FUT 3ms-stay.UN in home
 (If it rains, he will stay home)
- b. [ðað i-qqim g uxam]_Q [*xm* i-wwəθ unz^oar^o]_P
 FUT 3ms-stay.UN in home if 3ms-hit.PER rain
 (He will stay at home if it rains tomorrow)

In this example, the protasis is marked by the conjunction ‘*xm*’ (if) while the apodosis is not marked. The reason why the marked clause appears preferably before the unmarked one in many languages is probably to indicate that the situation described is hypothetical (see Comrie 1986 for a discussion of this point). However, since we were unable to consult a corpus of any variety of Berber, it is very difficult to decide whether this order is more frequent than the alternative one and, consequently, we cannot claim that it is the preferred order. Intuitively, the difference between (8a) and (8b) seems to be determined by contextual factors: (8a) seems to be primarily about the consequences of the possible rain while (8b) is more concerned with what the person referred to will do. To a certain extent, this difference is reflected in the corresponding English translations.

However, when the two clauses are marked, the only grammatical possibility is for the protasis to precede the apodosis. The following example illustrates this case:

9. [*mus* n-ffəy zik]_P [*uma* n-iwδ^o wasa]_Q
 If 1p-leave.UN early CON 1p-arrive.UN now
 (If we had left earlier, we would have arrived by now)

In this example, both P and Q are marked initially by a conjunction, thus indicating mutual dependency between the two clauses. That is to say, neither clause can stand on its own, unlike the previous case where the apodosis could serve as an independent clause. This holds also for conditionals with *'murid'*, in which the first clause can be verbless, as in the following example:

10. [murid i nta]_P [uma iwǝʔə-x zik]_Q
 If.not PART he, CON arrive.PER-1s early
 (Were it not for him, I would have arrived early)

The protasis in this example is nominal in the sense that the conditional conjunction and the pronoun (or a noun) are separated only by the particle *'i'* whose nature remains unclear. In any case, this particle cannot be of a verbal nature. Here, as in the previous example, the clause order cannot be reversed.

Alternatively, the two clauses may both be unmarked, though this can happen only under specific pragmatic conditions. But when this is the case, the protasis precedes the apodosis necessarily. The following examples are illustrative:

11. a. [aǝ i-zri wa]_P [a-θ i-ttabʃ wa]_Q
 PART 3ms-pass.UN this PART-him 3ms-follow.UN this
 (If/whenever one passes by, the other follows him)
 b. [a t-açrǝð lʃus]_P [a-ç bbi-x afus]_Q
 PART 2s-steal money, PART-you cut-1s hand
 (If you steal money, I will cut your hand)

In these sentences, the two clauses begin by the particle *'aǝ'*, which occurs with the unmarked stem of verbs. In some morpho-phonological contexts, this particle surfaces as *'a'*, as in (11b) and in the second clause of (11a). But this phenomenon is not of immediate relevance to the present discussion. Together, the imperfect aspect and the juxtaposition of the two clauses indicate that the proposition expressed by the first clause serves necessarily as the condition for the realization of that expressed in the second. It would appear that the lack of any marking on either clauses is the reason behind the ungrammaticality of the reversed order, given the risk of interpreting the consequent, occurring first, as the precedent. But the different marking of the two clauses, as in (9) and (10) above, does not pose such a risk; yet, the order is constrained. Therefore, an explanation must be sought elsewhere.

Concerning conditional markers, they vary considerably across Berber varieties. The two conjunctions, *'xm'* and *'mus'*, found in some of the previous examples, occur under different forms in different varieties, and it is possible that they show some internal variation as well. In the neighbouring dialects of the Middle Atlas, for instance, *'xm'* is realized as *'mš'* or as *'ʔiy'* and *'mus'* as *'muridis'* or *'murid'* or *'mur'*. In the Rif, far to the north, the first conjunction corresponds to *'mara'* and the second to *'imri'*. To the south, the equivalent of the first conjunction in Tashlhiyt is *'ʔiy'* and that of the second is *'mra'*. Distant varieties are likely to exhibit more

differences not only in form, but also in function. Unfortunately, these can only be encountered sporadically in the literature on Berber grammars, often in examples intended to illustrate some aspect of the language other than conditionals (but cf. Chafik 1991 and Kossmann 2013 for a preliminary list).

Markers of conditionality include, also, adverbials of time, place, and other constituents. These correspond to *Wh*-words attached to *-ever*, such as ‘*whenever*’, ‘*wherever*’, ‘*whoever*’ and ‘*whatever*’. The equivalents of these English adverbials in the Berber variety described here are ‘*milmi*’ or ‘*adday*’, ‘*manig*’, ‘*wənn*’ and ‘*may*’, respectively. Except for ‘*wənn*’, the other three conjunctions can be used as question words as well, as a comparison of the following examples illustrates:

12. a. *manig θəddam i θ'əħi?*
 where go.2p.PER last night
 (Where did you go last night?)
 b. [*manig θəddim*]_P [*aθ ddu-x ag-iðun*]_Q
 where go.2p.UN PART go.UN-1s with-you
 (Wherever you go, I will go with you)

The observation that question words and conditional conjunctions are often related cross-linguistically was first made by Traugott (1985). But it should be noted that some varieties of Berber, at least in Morocco, add the prefix ‘*ak*’ to these question words to signal conditionality (e.g., *akmilmi* ‘whenever’, *akwənn* ‘whoever’, etc.). This phenomenon is mentioned by Chafik (1991) but is not attested in the variety described here. This fact indicates that, although conditional conjunctions may originate from question words, they are very likely to develop in a different direction because of the semantic differences between conditionals and interrogatives. These markers of conditionality combine with verb aspect to express various degrees of hypotheticality to which we turn immediately.

Degrees of Hypotheticality

Hypotheticality is the most salient semantic component of conditionals. As explained in the first section, researchers converge on the idea that it forms a continuum rather than discrete categories, though languages differ as to the resources they provide to express different degrees of it. These resources relate to the speaker’s commitment to the occurrence or regularity of the situation expressed in the protasis as well as his attitude toward the likelihood of that expressed in the apodosis. These will be discussed jointly, below, in relation to Berber data.

The least degree of hypotheticality is found in what Athanasiadou and Dirven (1997) call “course of events” conditionals. This category can be illustrated by the following example:

13. [xm iyuða unbðu]_P [la tarw-n ulli iɕniwn]_Q
 If be.good.PER harvest PART bear.INT-3p sheep twins
 (If the harvest is abundant, the sheep give birth to twin lambs)

This sentence is not about a particular harvest, but rather about regular correlations between the two events expressed in the protasis and the apodosis. Therefore, the speaker commits himself to the idea that the two situations did occur and, more importantly, that they are likely to occur again. The use of the extensive particle '*la*' and the intensive form of the verb in the apodosis constitute the major resource to express this idea. The perfect aspect of the verb in the protasis is also essential to this category of conditionals; and its replacement by the imperfect aspect excludes the sentence from this category, or makes it completely ungrammatical. It should be noted that, because of the regularity of co-occurrence of the events expressed in the two clauses, the conjunction '*adday*' (when) is often preferred over '*xm*'. The difference between the two conjunctions seems to reside in the speaker's attitude toward the probable occurrence of the events in the future: '*xm*' apparently expresses a higher degree of hypotheticality than '*adday*'.

The most typical conditional structure in Berber, however, is that in which the verb in the consequent is in the unmarked form preceded by '*ɖaɖ*' or '*aɖ*'. An example of this structure is provided by (8a) above, repeated here as (14), with the addition of the adverbial '*asəkka*' (tomorrow):

14. [xm i-wwəθ unzʼarʼ asəkka]_P [ɖaɖ i-qqim g uxam]_Q
 if 3ms-hit.PER rain, tomorrow FUT 3ms-stay.UN in home
 (If it rains tomorrow, he will stay home)

In this example, the verb in the subordinate clause is in the perfect form while the verb in the main clause is in the unmarked form preceded by the future particle '*ɖaɖ*'. Both verb forms, however, refer to a future possibility, as indicated by the adverbial '*asəkka*'. This structure is fairly parallel to its English equivalent in that the speaker does not commit himself to the occurrence of P but thinks that Q is probable if P turns out to be true. But unlike the English future modal 'will', the particle '*ɖaɖ*' in the Berber variety described here may also have an epistemic use, as illustrated by (15):

15. ɖaɖ j-ili g uxam wasa
 FUT 3ms-be.UN in house now
 (He must be at home now)

The presence of the adverbial '*was*' (now) in (15a) leaves no possibility for a future reading of the sentence; therefore, it must be a deduction, expressed mainly by the particle '*ɖaɖ*', usually associated with the future time. On the basis of this remark, Q in (14) could be a prediction, but it could also be a deduction based on past correlations between P and Q. Under this second reading, (14) would be a "course

of events” conditional and, consequently, has a lesser degree of hypotheticality than the first reading.

Another grammatical possibility to reduce the degree of hypotheticality (and increase the probability) of a conditional is to use the particle ‘*að*’ instead of ‘*ðað*’ in the apodosis. (16) illustrates this point:

16. [xm θ-tt-waçər lçsib-θ in-u]_P [að zznz-x alim]_Q
 if 3fs-PAS-steal livestock-f of-me PART sell-1s hay
 (If my livestock are stolen, I will sell the hay)

Although the difference between this example and (14) involves only the nature of the particle used in the apodosis, it seems that Q in (16) is believed by the speaker to be more probable than that in (14). The particle ‘*að*’ apparently expresses a certain degree of the speaker’s control on the course of events that ‘*ðað*’ does not. Indeed, if used in (14), ‘*að*’ would imply that the speaker has control on the individual referred to in the apodosis.

Hypotheticality seems to be affected by the aspect of the subordinate verb. Although the verb in the protasis is usually in the perfect in Berber, it may also occur in the imperfect. The following example is a different version of (14) that illustrates this fact:

17. [xm ðað i-wwəθ unzʼarʼ asəkka]_P [að i-qqim g uxam]_Q
 if FUT 3ms-hit.PER rain, tomorrow PART 3ms-stay.UN in home
 (If it rains tomorrow, he will stay home)

In this example, P implies an acceptance by the speaker of the assumption, probably made initially by the hearer, that it is very likely it will rain the following day. In other words, P is entertained as more than just an imaginative potentiality, but rather as an event of highly probable occurrence. On this basis, the speaker also believes that, for him to stay at home is the most likely, or even the most appropriate, decision to make. As explained in the previous paragraph, this prediction can be based partly on the speaker’s authority over the referent. The use of ‘*ðað*’, instead of ‘*að*’, would cancel out this implicature. This nuance of the meaning seems to be limited to varieties such as Zemmour Tamazight that distinguish between the two particles.

Another resource to express a lesser degree of probability is the conditional conjunction itself. So far, most of the examples cited of Berber conditionals use ‘*xm*’, but of equal importance is the conjunction ‘*mus*’, as in this example:

18. [mus-d j-iwəðʼ wasa]_P [uma n-dda]_Q
 If-here 3ms-arrive.PER, now PART 1p-go.UN
 (If he arrived now, we would leave (immediately))

It should be noted that, although the verbs in both clauses are in the perfect form, they are used here with a present time reference, as indicated by the adverbial ‘*was*’. Apparently, this use implies a lesser probability of occurrence, or regret of the

non-occurrence of the event. The use of the particle ‘*að*’ is also possible here and seems to imply that immediate leaving is still possible, though its probability is low—expressed mainly by the conditional conjunction.

‘*mus*’ can also be used with future situations just like ‘*xm*’. Thus, (18) could be turned into (19):

19. [mus ðað j-awð^ε asəkka]_P [uma gguni-x-θ]_Q
 If FUT 3ms-arrive tomorrow, PART wait.PER-1s-him
 (If he arrives tomorrow, I would wait for him)

In this example the conditional conjunction could be replaced by ‘*xm*’ if the necessary changes are made. These changes include the deletion of ‘*uma*’ from the apodosis and the replacement of the perfect form of the verb ‘*gguni*’ by its imperfect form, preceded by the particle ‘*að*’. The difference between ‘*mus*’ and ‘*xm*’ would be one of probability: (19) certainly expresses more doubt about the possible arrival of the person tomorrow than would be the case if ‘*xm*’ had been used instead. It seems that the perfect form of the verb in the apodosis plays no particular role in this reading since it can be replaced by the imperfect form ‘*a-θ ggani-x*’ without increasing, or decreasing the hypotheticality of the event.

The argument that ‘*mus*’ expresses a higher degree of hypotheticality than ‘*xm*’ is made clearer with past situations. The examples (20) provide a good comparison:

20. a. [xm i-qləʃ inijn-wassnnat]_P [ðað j-ili j-iwð^ε ið^ʕHi]_Q
 if 3ms-depart.PER before-yesterday, FUT 3ms-be.UN 3ms-arrive last night
 (If he had departed the day before yesterday, he must have arrived last night)
 b. [mus i-qləʃ inijn-wassnnat]_P [uma j-iwð^ε ið^ʕHi]_Q
 if 3ms-depart.PER before-yesterday, FUT 3ms-be.UN 3ms-arrive last night
 (If he had departed the day before yesterday, he would have arrived last night)

The time adverbials in the two sentences clearly indicate that the situations described, both in the protasis and the apodosis, have a past time reference. But while the speaker assumes that P is true in (20a) and makes a deduction in Q on the basis of that assumption, in (20b) he asserts that P is false and, consequently, Q must also be false. In more precise terms, while (20a) is hypothetical, (20b) is counterfactual, as their respective English translations indicate. Although the aspect used in the apodosis is different in (20a) and (20b), it could be made similar without affecting the readings of the two sentences. This fact suggests that the semantic difference is essentially due to the difference in conjunctions.

Another conjunction that is usually associated with the counterfactual reading is ‘*murid*’ illustrated by (10) above. The example is repeated here as (21):

21. [murid i nta]_P [uma iwð^εə-x zik]_Q
 If.not PART he, CON arrive.PER-1s early
 (Were it not for him, I would have arrived early)

(21) has no time adverbials. Besides, the protasis lacks any verbal form. Yet, the sentence has only a counterfactual reading. Dialects which do not mark the apodosis also show the same reading of the corresponding construction, as testified by this example, borrowed from Chafik (1991):

22. [murid i fəgg]P [i-nʔa-θn]Q
 If.not PART you.ms 3ms.kill.PER-them
 (If it were not for you, he would have killed them)

In both (21) and (22), the verb in the apodosis is in the perfect form, but it could be replaced by an imperfect form (with the particle ‘*að*’) without affecting the counterfactual reading. As concluded in the preceding paragraph, this reading can only be due to the conjunction.

The conjunctions ‘*murid*’ and ‘*mus*’ are apparently related. The dialectal variant ‘*muridis*’ is probably the full form from which the two conjunctions are diachronically derived. Since ‘*murid*’ encapsulates the element of negation, much as ‘*unless*’ does in English, it is possible that it is morphologically complex. In more precise terms, ‘*ur*’ could be the same negative verbal prefix attested in most Berber varieties, if not all of them (cf. Taifi 1993). The ‘*is*’ part of ‘*muridis*’ is also reminiscent of a similar conjunction usually used with yes-no questions. This should come as no surprise given that conditional conjunctions are often derived from question words, as is testified by some uses of ‘*if*’ in English (cf. Traugott 1985). On the basis of these conjectures, a conditional marker with an initial ‘*m*’ could be reconstructed for an older stage of the language. As noted earlier, the conditional conjunction in some varieties of Berber, such as Tarifit, is ‘*mara*’ (compare also ‘*ma*’ in Taqbaylit). Accordingly, we might advance that ‘*mus*’ is originally constituted of a conditional marker and a question word. If this reconstruction turns out to be supported, we would have accounted for the semantic similarity between the conjunctions ‘*murid*’ and ‘*mus*’.

On the basis of the data described in this section, two major conclusions could be drawn. The first concerns the categorization of conditionals according to conjunction while the second relates to the role of aspect. Given the semantic difference between ‘*xm*’ and ‘*mus*’, it appears that Berber distinguishes between two types of conditionals: (a) those which are judged by the speaker to be possible and are characterized by the use of ‘*xm*’, and (b) those which are judged to be unlikely, distinguished by ‘*mus*’ or ‘*murid*’. As evidenced by the abovementioned examples (18 and 19), ‘*mus*’ is not limited to counterfactuals and can also be used with hypothetical events. With both hypotheticals and counterfactuals, however, it expresses the doubtful attitude of the speaker. Similarly, ‘*xm*’ is not limited to present or future hypothetical events and can also be used with past hypotheticals, as shown above by (20a). Therefore, the difference between the two conjunctions is essentially one between ‘hopeful’ hypotheticals and ‘doubtful’ ones. Counterfactuals are not treated as distinct from other high hypotheticality conditionals, though they may serve as their prototype. This implies that Berber distinguishes between two basic categories of conditionals on the basis of their degree of hypotheticality. Since the two are marked by different conjunctions, they could be considered as discrete categories

rather than points in the same cline. Within each category, however, linguistic resources other than conjunctions can be deployed to express various degrees of probability. The next section discusses some peripheral types of conditionals.

Some Related Adverbial Constructions

Among the types of adverbial clauses that overlap with conditionals are concessives, and wishes. The overlap could be semantic only—but it could also be morpho-syntactic.

We will first begin with time concessives. Concessive conditionals are often similar to ordinary, or prototypical, conditionals, but as Thompson et al. (2007: 261) note, “conditional concessive clauses do carry additional presuppositions not signaled by ordinary conditionals”. In English, these presuppositions are usually expressed by the first element of the compound ‘*even if*’. The following example provides a comparison between Berber concessives and their English equivalents:

23. a. waxxa i-bð'a unẓ'ar⁶ assnnat⁶, fɣə-n að urar-n
 although 3ms-fall.PER rain yesterday, exit.PER-3p PART play.UN-3p
 (Although it rained yesterday, they went out to play)
 b. [waxxa i-bð'a unẓ'ar⁶ asəkka]_B [ðað fɣə-n að urar-n]_Q
 although 3ms-fall.PER rain tomorrow, FUT exit.UN-3p PART play.UN.3p
 (Even if it rains tomorrow, they will go out to play)

(23a) is a simple concessive in contrast to (23b), which is a concessive conditional. The difference between the two is that the first presupposes the truth of the propositions expressed by both clauses, whereas the second makes no such presupposition. However, the fact that both types use the same conjunction suggests that conditionals and concessives are connected. Semantically, (23b) asserts Q and presupposes (a) that there is an expectation that the proposition “if it rains, they will go out to play” would not be true and (b) that there is a belief that the proposition “if it doesn’t rain, they will go out to play” is likely (cf. Thompson et al. (2007: 262). Köng (1986) speaks of the distinction between concessives and conditional concessives. In comparison, ordinary conditionals do not carry either presupposition although, like concessives, they assert Q.

Related to conditionals are also wishes. In English, wishes could be expressed by “*I wish*” or by the subordinator “*if only*”. The use of the conditional ‘*if*’ is suggestive of the connection with conditionals. In Berber, too, the same conjunction is found in wishes and some conditionals, as the following example testifies:

24. mus ɣas i-bð'a unẓ'ar⁶
 if only 3ms-fall.PER rain
 (If only it (had) rained)

Although the verb is in the perfect form, without a time adverbial this sentence could be a wish about the past, the present, or the future. This fact proves that aspect does

not play any crucial role in determining time reference of conditionals and related constructions such as wishes—an argument advanced earlier. The use of ‘*mus*’ also supports an earlier argument according to which this conjunction expresses what we called “doubtful” hypotheticality. In other words, the situations referred to in the ‘*mus yas*’ construction are conceptualized as very unlikely irrespective of whether the attitude is one of regret about a past situation, a complaint about a present one, or a wish for a future one. This might seem surprising given that a wish, unlike regret about the past, might be fulfilled no matter how distant it may be. English, for example, does distinguish between the two types by using different tense and aspect markers (viz. simple past for wishes and past perfect for regrets). In contrast, Berber does not formally mark this dichotomy, though the difference between the two would certainly be communicated by contextual clues, that is to say, pragmatically.

On the basis of the conjunction used, wishes are more closely related to conditionals than concessives. This formal relationship between conditionals and wishes seems to reflect a semantic connection that is weaker than that between conditionals and concessives. The connection, apparently, concerns the unreality of the apodosis: in conditionals, it is hypothetical, just like the proposition expressed in wishes, whereas in concessives its truth is asserted, as explained above.

Conclusion

This paper investigated conditional constructions in Berber, specifically the variety spoken in the Zemmour plateau and the Middle Atlas Mountains in Morocco. The main points made are the following: (a) with some conjunctions, only the protasis is marked, while with others, both clauses are marked; (b) when only the protasis is marked, the order of clauses is relatively free, but when the apodosis is also marked, the order is fixed, with the protasis coming first; (c) when neither clause is marked the order is also fixed, as before; (d) constructions in which both clauses are marked express a high degree of hypotheticality than those in which only the protasis is marked; (e) counterfactuality is not formally distinct from high hypotheticality; (f) aspect may be exploited to express various degrees of probability, but conjunctions play a bigger role; (g) within marginal types, wishes are more closely related to conditionals than concessives. However, given the variation, both between and within Berber languages, it is likely that these conclusions may not hold for distant varieties. Therefore, further investigation of conditional constructions in other varieties of the language is needed for a better understanding.

References

- Athanassiadou A, Dirven R (1995) Typology of if-clauses. In: Casard E (ed) *Cognitive linguistics in the redwoods*. Mouton, Berlin, pp 609–654
- Athanassiadou A, Dirven R (1997) Conditionality, hypotheticality and counterfactuality. In: Athanassiadou A, Dirven R (eds) *On conditionals again*. John Benjamins, Amsterdam, pp 61–96

- Basset A (1929) *La langue berbère. Morphologie. Le verbe. Étude de thèmes*. E. Leroux, Paris
- Basset A (1952) *La Langue Berbère*. Oxford University Press, Oxford
- Bouhkris F, Boumalk A, El-Moujahid E, Souifi H (2008) *La Nouvelle Grammaire de l'Amazighe*. Publications de l'Institut Royal de la Culture Amazighe, Rabat
- Chafik M (1991) *ʔarbaʕ wa ʔarbaʕūna darsan fī alluḡati al-ʔamāzīḡiyah*. Alnnašr alʕarabī alʔifrīqī, Rabat
- Chaker S (1989) Aspect. In: Camps G (ed) *Encyclopédie berbère*. Editions Peeters, Paris
- Comrie B (1986) Conditionals: a typology. In: Traugott E, ter Meulen A, Relly J, Ferguson C (eds) *On conditionals*. Cambridge University Press, Cambridge, pp 77–99
- Destaing E (1920) *Etude sur le Dialecte des Aït Seghrouchen*, (Moyen Atlas Maroain). Leroux, Paris
- König E (1986) Conditionals, concessive conditionals and concessives: areas of contrast, overlap and neutralization. In: Traugott E, ter Meulen A, Relly J, Ferguson C (eds) *On conditionals*. Cambridge University Press, Cambridge, pp 229–246
- Kossmann MG (2013) *The Arabic influence on Northern Berber*. Brill, Leiden
- Laoust E (1924) *Cours de Berbère Marocain: Dialecte du Maroc Central*. Geuthner, Paris
- Nicolle S (2017) Conditional constructions in African languages. *Studies in African Linguistics* 46(1–2):1–15
- Penchoen TG (1973) *Tamzight of the Ayt Ndhir*. UNDENA Publications, Malibu
- Quitout M (1997) *Grammaire Berbère: rifain, tamazight, chleuh, kabyle*. L'Harmatton, Paris
- Sadiqi F, Ennaji M (2004) *A grammar of Amazigh*. Publications of the Faculty of Letters (Pars Lettres 25), Fes
- Taifi M (1993) L'hypothèse en berbère. In: Drouin J, Roth A (eds) *A la croisée des études libyco-berbères. Mélanges offerts à Paulette Galand-Pernet et Lionel Galand*. Geuthner, Paris, pp 215–227
- Thompson S, Longacre R, Hwang SJ (2007) Adverbial clauses. In: Shopen T (ed) *Language typology and syntactic description*. Cambridge University Press, Cambridge, pp 237–300
- Traugott E (1985) Conditional markers. In: Haiman J (ed) *Iconicity in syntax*. John Benjamins, Amsterdam, pp 289–307



Frequency Norms in Tashlhiyt: A Pilot Study 21

John Alderete, Jane Li, and Rachid Ridouane

Contents

Introduction	444
Methods	445
Word Frequency	446
Consonant and Vowel Frequency	446
Conclusion	451
Appendix	453
References	453

Abstract

This article gives a preliminary account of lexical and sub-lexical frequency norms in Tashlhiyt, an Amazigh language of Morocco. A corpus of nine texts with roughly 19,000 words was analyzed as a list of structured representations for words that distinguish part of speech, open/closed class, morphological makeup, and surface versus lexical representations. A lexicon was built from the corpus to identify unique word types and to document the frequencies of these types. The frequencies of sub-lexical phonological segments are also documented and we investigate the impact of phonological representation, morpho-syntactic categories, and the type/token distinction on segment frequency. All of the frequency norms are available as open data sets.

J. Alderete (✉)

Simon Fraser University, Burnaby, BC, Canada
e-mail: alderete@sfu.ca

J. Li

Johns Hopkins University, Baltimore, MD, USA
e-mail: sli213@jhu.edu

R. Ridouane

Laboratoire de Phonétique et Phonologie, CNRS & Sorbonne Nouvelle, Paris, France
e-mail: rachid.ridouane@sorbonne-nouvelle.fr

Keywords

Frequency norms · Word frequency · Segment frequency · Corpus linguistics · Tashlhiyt · Amazigh

Introduction

Frequency norms are critical to psycholinguistic research, natural language processing applications, corpus studies, and, increasingly, in linguistic analysis (Bod et al. 2003; Gries 2016). For example, our understanding of language comprehension and production in Indo-European languages like English and Dutch depends on well-constructed and validated data sets (Baayen et al. 1996; Kessler and Treiman 1997). For a variety of reasons, the creation of high-quality data sets has been largely focused on a small number of majority languages with large numbers of speakers and representing socio-economic power. For example, Anand et al. (2011) reports that among the 550 corpora available from the Linguistic Data Consortium, just five languages account for 85% of the data sources. The goal of this paper is to address this gap by creating a core set of frequency norms for word and segmental patterns in Tashlhiyt, an Amazigh language of Morocco.

Past research on Tashlhiyt has documented the frequencies of selected phonological patterns. In particular, quantitative accounts have been given of phonotactic patterns (Boukous 1987), consonant co-occurrence patterns (Bensoukas 2014), and contact phenomena (Boukous 2012), but a general account of segment frequencies and segment combinations has not yet been produced. As a sign of the interest in this topic, Orzechowska and Ridouane (2018) have recently given a detailed quantitative account of the patterns in long strings of consonants in vowel-less verbal roots. Their investigation of these combinations using principal component analysis identified the key features shaping these sequences and provided important insight into the phonotactic constraints on these verbs. However, a comprehensive account of segment frequencies in other word classes, word frequencies, and the differential impact of type and token frequency has yet to be developed.

In this article, we document segment and word frequency in a corpus of nine texts of Tashlhiyt (Boukous 1977). The corpus has 18,827 word tokens of 4195 lexical items from all word classes. As the first general account of these frequencies, we focus here on a few foundational questions and problems. After describing our methods (section “[Methods](#)”), we document word frequency in the lexicon and investigate high-frequency words (section “[Word frequency](#)”). In section “[Consonant and vowel frequency](#)”, we document the frequencies of consonants and vowels, and specifically examine the impact of word class (i.e., noun versus verb), the open class/closed class distinction, type versus token frequency, and the level of the phonological representation (i.e., surface versus lexical). Section “[Conclusion](#)” summarizes these results and sketches how frequency norms can be developed in future studies.

Methods

The frequencies of words and sounds are drawn from Jebbour, et al. (2021), an electronic corpus of nine texts from Boukous (1977). The transcript of each text was digitized and analyzed as a list of structured representations, or *n*-tuples. Each word was encoded as a 6-tuple that analyzes different aspects of the word. That is, the word was ordered in the list as it appears in the text and associated with specific information, as shown in Table 20.1 for the sentence *Imma !!bisara mnfk at tga i kilu* ‘and how much does the bean soup cost?’

The first two elements of the 6-tuple give the order in the text and the title of the text. The third (SOUND) and fourth (MORPHEMES) give the surface and lexical representation, respectively. The surface form is the phonetic form of a word as it is pronounced, and the lexical representation is the surface form before any phonological rules have been applied. The M-GLOSS (fifth element) gives a gloss of the morpheme and its grammatical functions (see the documentation of Jebbour et al. (2021) for the coding principles of grammatical morphemes), and CATEGORY specifies N for nouns, V for verbs, and \emptyset for all other word classes. The first word of every sentence is preceded by a 6-tuple containing ‘New_Sentence’, indicating the end of the previous sentence and the beginning of a new one.

The sound frequencies documented below were established with searches on different components of the 6-tuples. For example, counts of surface segments in nouns were made by searching in the SOUND column and only in words with ‘N’ under CATEGORY. In this way, we were able to distinguish lexical versus surface phonological representations, nouns versus verbs, and content versus functional items (i.e., all words that are not nouns or verbs), and so forth.

We also distinguished type versus token frequencies in the following way: token frequencies are simply counts of all words in the corpus. Word types are unique words with a distinct morphological makeup. Concretely, items in the dataset with the same morphological makeup (underlying form), gloss, and part of speech are treated as one type. This entails that items with differing surface forms can be joined

Table 20.1 Illustration of structured representations of words

Order	Texte	SOUND	MORPHEMES	M-GLOSS	CATEGORY
...					
443	lxdmt	New_Sentence	New_Sentence	New_Sentence	
444	lxdmt	imma	[imma]	[whereas]	
445	lxdmt	!!bisara	[!-!bisara]	[bean soup]	N
446	lxdmt	mnfk	[mnfk]	[how much]	
447	lxdmt	at	[ad]	[AD:COMP]	
448	lxdmt	tga	[t-ga]	[3FS-cost:pf]	V
449	lxdmt	i	[i]	[for]	
450	lxdmt	kilu	[kilu]	[one kilogram]	N
451	lxdmt	New_Sentence	New_Sentence	New_Sentence	

as one type. In such contexts, following common practice, we chose the most frequent surface variant for sound-level type tabulations (Connine, et al. 2008; Pitt, et al. 2011). The word /a-mɰar/ ‘community leader’, for example, has two surface forms [jamɰar] and [amɰar]. Since [amɰar] is the more frequent variant, our type frequencies reflect this by adding two counts of [a], one count of [ɰ], [r], and [m], respectively. We are aware that some studies of morphologically rich languages also specify root type based on shared roots and not shared grammatical morphemes (Aksan and Yaldir 2012; Gerz, et al. 2018), but we leave typing based on roots for future research (see below). A lexicon of word types was built for the entire corpus, and type frequencies were generated using this lexicon.

Two large data tables documenting the frequency norms and a set of scripts for extracting frequencies are available from the first author’s website. The contents of data tables for words (`wordfreq_master`) and sounds (`soundfreq_master`) are explained in more detail in the Appendix.

Word Frequency

The corpus contains 18,827 word tokens and, as explained in the methods above, these tokens represent 4915 word types with distinct form or morphological makeup. The word frequency of specific lexical items, among other things, is given in the `wordfreq_master` data table (see Appendix), but we report here on the salient facts of the corpus as a whole. In general, the vast majority of words have very low frequency: 3191 items (65%) occur only once in the corpus, and another 660 (13%) occur just twice. Furthermore, there are just 18 lexical items with a frequency of 100 or greater, listed in Table 20.2. As one might expect, these are dominated by grammatical morphemes: only two forms of *say* and two names that are common in these texts rank among these other high-frequency function words.

The relative frequency of nouns, verbs, and closed class morphemes are reported in Table 20.3. While nouns and verbs together as a group have about the same frequency in the corpus as all other words, they greatly outnumber these words in the lexicon, which is expected, because these other words are, by and large, closed class items with higher corpus frequency.

The general skewing of words towards the low end of the frequency spectrum is shown in Fig. 20.1, as also the Zipfian-like distribution in the lexicon (Zipf 1949). There are a relatively small number of very high-frequency words, and word frequency falls very quickly (from left to right in the graph) until it plateaus at the low end of the spectrum.

Consonant and Vowel Frequency

We now report the segment frequencies for all segments, including consonants and vowels in a variety of conditions of interest. We begin with frequency distributions for consonants (Table 20.4) in surface representations of all words. The counts of the

Table 20.2 All lexical items with frequency greater than 100

Lexical item	Part of speech	M-Gloss	Frequency	Probability
ad		AD:Comp	518	0.02842
ias		to-dat-3S	423	0.02321
a		Voc	301	0.01651
ar		AR	286	0.01569
ur		neg	270	0.01481
i		to	224	0.01229
as		dat-3S	206	0.01130
inna	V	3MS-say:pf	193	0.01059
iini	V	3MS-say:ao	187	0.01026
awa		so	168	0.00922
d		and	138	0.00757
ns		PossDet:3S	138	0.00757
iji		dat:1S	138	0.00757
gi-s		loc:3MS	120	0.00658
ḥasan	Name	Hassan	118	0.00647
n		of	113	0.00620
!rbbi	Name	God	110	0.00604
nkki		1S:me	104	0.00571

Table 20.3 Tokens, types, and mean frequency by part of speech category

	Tokens	Types	Mean frequency
Nouns	4363	1490	2.9282
Verbs	4567	2169	2.1056
Other	9297	1256	7.40207

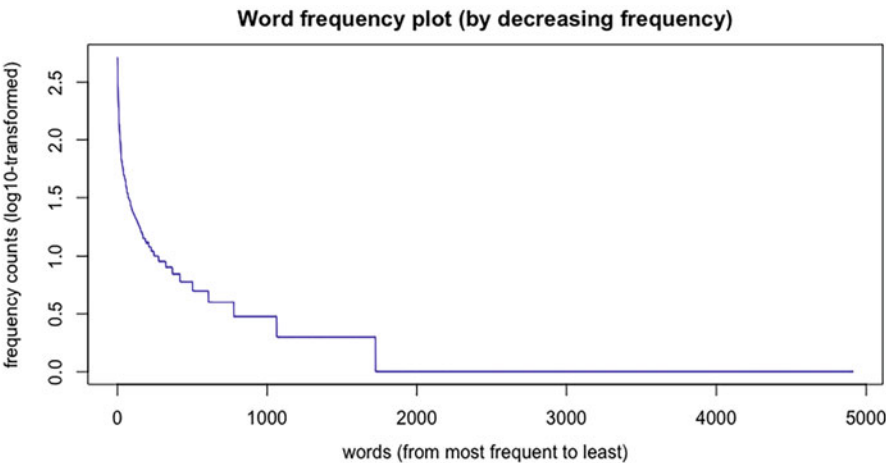


Fig. 20.1 Distribution of word frequency in the corpus

Table 20.4 Consonant frequencies, all words, surface representations (color coding for top six segments)

Singleton	Token	Type	Geminate	Token	Type
b	822	391	bb	237	55
t	3935	1608	tt	637	334
d	1979	643	dd	431	140
t ^ɕ	155	106	tt ^ɕ	189	92
d ^ɕ	328	194	dd ^ɕ	34	24
k	1617	570	kk	395	153
g	977	295	gg	103	68
k ^w	3	3	kk ^w	1	1
g ^w	2	2	gg ^w	0	0
q	179	101	qq	312	197
q ^w	1	1	qq ^w	0	0
f	894	404	ff	73	39
s	2831	705	ss	427	185
z	447	197	zz	183	97
s ^ɕ	156	105	ss ^ɕ	77	49
z ^ɕ	169	106	zz ^ɕ	19	16
ʃ	385	204	ʃʃ	100	62
ʒ	151	96	ʒʒ	77	42
ʃ ^ɕ	27	21	ʃʃ ^ɕ	28	14
ʒ ^ɕ	25	15	ʒʒ ^ɕ	49	9
χ	320	156	χχ	47	14
ʁ	1819	642	ʁʁ	1	1
χ ^w	0	0	χχ ^w	0	0
ʁ ^w	4	3	ʁʁ ^w	0	0
h	769	286	hh	7	4
ɦ	685	191	ɦɦ	4	3
ʕ	562	296	ʕʕ	0	0
m	2214	866	mm	293	115
n	3859	1365	nn	937	265
n ^ɕ	133	76	nn ^ɕ	12	10
l	2340	966	ll	990	315
l ^ɕ	152	88	ll ^ɕ	122	33
r	2737	818	rr	132	65
r ^ɕ	660	333	rr ^ɕ	70	43
w	1837	650	ww	24	16

six most frequent segments are color-coded (from highest to lowest: yellow, green, aqua, magenta, grey, and red). Within this set of high-frequency segments, *t*, *s*, *n*, and *l* rank in the top six in both token and type frequencies and in singleton and geminate sounds. The relative ranks, however, do differ in these classes. For example, *kk* and *qq* have a higher relative frequency in geminates than in singletons. Furthermore, while token and type frequency tend to be correlated, there are clear differences. For example, the relative rank of *s* drops from token to type frequency, but the opposite pattern is found for *l*. This table also establishes quantitatively two intuitions that many Amazigh linguists have for segment frequencies, namely that geminates and emphatics are far less frequent than their singleton and non-emphatic counterparts. Finally, some segments are exceedingly rare or non-existent in this corpus, including

the velars with secondary labialization, the voiced geminate uvular *ʙʙ*, and the geminate pharyngeal *ʕʕ*.

These frequencies are aggregated into manner, place, and voicing classes in Table 20.5, which excludes glides because of their vowel-like status. Many of these frequencies are expected. For example, coronals have the highest token and type frequency among the place classes consistent with the fact that they dominate the inventory of 36 consonants. Interestingly, however, the frequencies of labials and dorsals are comparable, but there are roughly three times as many dorsals as labials in Tashlhiyt. Likewise, voiced consonants are both more numerous in the inventory and higher in frequency. Fricatives, however, outnumber stops in the inventory by a proportion of roughly 3–2, but stops are much more frequent than fricatives in both type and token frequencies. The descriptive accounts of these natural classes, and many others, can be explored further with the data supplement.

As for vowel frequencies (Table 20.6), this trio of segments decreases in frequency, *a* > *i* > *u*, in both token and type frequencies.

As far as how vowels pattern with the rest of the inventory, the percentage occurrence of vowels relative to consonants is roughly 35%/65% in tokens and 30%/70% in types.

As explained in section “Methods”, the structured representations in our corpus distinguish lexical and surface representations of words. How well are the frequencies in these representations correlated? Here, and throughout, we assess this by giving the correlation coefficients for consonants (including geminates) and all segments (all consonants and vowels). Thus, the frequency distribution of all consonants is a vector of 72 segment counts and 75 counts for all segments. To compare two vectors, we assessed the relationship between two classes

Table 20.5 Frequencies of natural classes

Type	Natural class	Token	Type
Place	Labial	4533	1870
	Coronal	24,983	9441
	Dorsal	5781	2398
	Pharyngeal	1338	586
	Glottal	898	407
Manner	Stop	12,337	4978
	Fricative	10,336	3962
	Nasal	7448	2697
	Liquid	7203	2661
Voicing	Voiced	23,759	8884
	Voiceless	13,565	5414

Table 20.6 Vowel frequencies, all words, surface representations

	Token	Type
i	8175	2645
u	2488	868
a	11,907	3270

(e.g., frequencies in nouns versus verbs) with Pearson’s correlation coefficient r . As shown in Table 20.7, the correlation between frequencies in lexical and surface representations is very strong in both token and type frequency. Also, in both token and type frequencies, these correlations improve with all segments, presumably because of the large counts of vowels. In sum, one can say that the representations differ slightly before and after phonological processes have been applied, but not by much.

Though the differences between lexical and surface forms are minor, we, nonetheless, analyzed the mismatches in case they might be relevant to a future study. Of the 4915 items in the lexicon, only 503 (10.23%) mismatch with the surface form; though these mismatches do not include changes in surface form due to emphatic spread, which is substantial. The average string edit distance of the mismatches is 1.14. Therefore, while a non-trivial percentage of the lexicon has a mismatch, the mismatches are typically in a single segment, consistent with the high correlations in Table 20.7. Among the 503 mismatched words, roughly one-fifth involve a deletion or addition, and the remaining phonological patterns involve well-known alternations in Tashlhiyt, including vowel-glide alternations and devoicing, as broken down in Table 20.8.

Next, we ask if and how segment frequencies are affected by morpho-syntax. It is well known that the distribution of segments are different in content and function words, often with function words having more restricted sound inventories (Beckman 1999; Willerman 1994). Furthermore, we expect large differences in segment distributions because of Tashlhiyt morphology. For example, approximately 38% of verbs in our corpus start with $i \sim j$ due in part to the prevalence of the third person masculine prefix, whereas only 11% of nouns start with these sounds, and verbs also introduce schematic patterns not found in nouns, like the occurrence of geminates in imperfectives. Table 20.9 gives the correlations between nouns and verbs, and content and functional items, for both token and type

Table 20.7 Correlations between surface and lexical frequencies

	Token	Type
Consonants	0.9856	0.9952
All segments	0.9935	0.9974

Table 20.8 Phonological patterns in mismatched words

Pattern	Example	N (of 503)
Changes segment count		(n=92)
Deletions	/ra-nt/ → ran	56
Additions	/ajt/ → jajt	36
No changes segment count		(n=411)
$i \sim j$	/i-ili/ → jili	271
$t \sim d$	/ard/ → art	38
$u \sim w$	/uarjal/ → warjal	31
$ss \sim zz$	/issnza/ → izzna	21
Other	/mad/ → maj	50

Table 20.9 Correlations based on morpho-syntactic structure

	Token		Type	
	Consonants	All segments	Consonants	All segments
Noun vs. Verb	0.7917	0.8346	0.8503	0.8810
Content vs. Functional	0.7673	0.9021	0.8589	0.9373

Table 20.10 Correlations between token and type frequencies

	Consonants	All segments
Nouns	0.9873	0.9952
Verbs	0.9831	0.9815
Function words	0.9188	0.9785
All words	0.9748	0.9873

frequencies. These results show that, while correlated, there are significant differences between nouns and verbs on the one hand and content and functional items, on the other. The correlations in consonant and segment frequencies are also stronger in type frequencies, as one would expect because the impact of high-frequency words (skewed toward functional items) is reduced.

Finally, are token frequencies correlated with type frequencies? That is, are the frequencies as they are reflected in the entire corpus well-correlated with the frequencies in the lexicon? Table 20.10 breaks these frequencies down by the four morpho-syntactic classes, and the distinction between consonants versus all segments. In general, token and type frequencies are highly correlated, and they become stronger when the entire set of segments is included, as observed above. Nouns and verbs do not seem to differ very much, except with the slightly weaker correlation in verbs for all segments. Perhaps the most noticeable difference is in function words, which have a .07 drop in r in consonants relative to nouns and verbs. We expect this because a higher percentage of very high-frequency words are function words. Thus, when the effect of these words is lost in the lexicon, we expect changes in the frequency distribution of the sounds of these words.

Conclusion

This report gives a first approximation of the frequency distribution of words and segments in Tashlhiyt. We have examined a relatively large corpus of words, built a lexicon for this corpus, and analyzed the impact of the token/type distinction, lexical and surface representation, and morpho-syntactic classes on frequency norms. It turns out that the distribution of segments is affected only slightly by the lexical-surface and token-type distinctions, though specific segments, sometimes do exhibit non-trivial differences. Distributions based on different morpho-syntactic classes (i.e., noun versus verb and content word versus functional item) are also correlated, but the correlations are much weaker. The morphology of these words and the basic distinction between open and closed class items, therefore, affect the relative frequencies of sounds.

These results, and the data they are built from, can be employed in a variety of ways. They can be used in psycholinguistic experiments where word and segment frequency need to be balanced or manipulated. For example, one can use the word frequency data supplement to find stimuli from high- versus low-frequency classes. Likewise, experimental items can be balanced for type frequency by using the frequency information from the sub-lexical data supplement. In addition, all of these frequency calculations can be tailored to the experimental task, selecting frequencies from a particular kind of representation (lexical or surface), morpho-syntactic context (noun, verb, functional item), and frequency type (token or type). Beyond these applications, the frequency distributions documented here can be used by linguistic analyses that require frequency to determine core ingredients of phonological analysis (Blevins 2004; Bybee 2001; Hayes and Wilson 2008).

While we believe the frequency norms documented here are useful, this report is just the beginning of documenting frequency norms. In particular, we envision several ways of improving them so that they can better inform other kinds of investigations. First, our lexicon uses a word-based typing system, but perhaps a better way to investigate the token/type distinction is by recognizing root-based types. Though Tashlhiyt does not have distinct inflectional classes, it has a rich verb and noun morphology, and so it makes sense to distinguish word types on the basis of a shared root or stem, as advocated by Aksan and Yaldir (2012) for Turkish. Doing so will reduce the number of word types and, as a result, it will change the type frequencies of segments and word frequencies. We do not think this will affect segment frequencies very much, but it will affect word frequencies, as a considerable amount of word types in our current lexicon will be collapsed into a single type.

Another way to further develop the project is to create a syllabified surface representation in the corpus and use it to derive frequencies for whole syllables and segments cross-classified by syllabic role (e.g., *n* in an onset as opposed to acting as a syllable nucleus). The algorithms for syllabifying segmental strings are well-known (Dell and Elmedlaoui 2002; Jebbour 1996), and so enriching the structured representations in this way is a tractable problem. Furthermore, Tashlhiyt is typologically unusual in that it allows all consonants, including non-sonorants, to occupy the syllable nucleus. An understanding of the frequencies of these different syllable nuclei and the frequencies of syllables containing them is of potential interest to investigations that seek to understand how syllable composition affects language behavior. For example, how does the syllable inventory of Tashlhiyt differ from well-known Indo-European languages, and are the syllables that are unattested in these languages, but attested in Tashlhiyt, spoken and perceived in different ways?

Finally, while the frequency norms we document here can inform new investigations, they have not yet been validated in the sense that they have been shown to predict certain well-known linguistic behaviors. For example, we expect high-frequency words to be classified as words of the language by native speakers faster than low-frequency words in lexical-decision tasks, and we expect words with high-frequency syllables to be spoken faster than other words (Levelt and Wheeldon 1994; Whaley 1978). In terms of psycholinguistic behavior, Tashlhiyt is an understudied language, and so it is difficult to find relevant behavioral data to correlate

with the frequency distributions we have established here. There is no substitute to collecting behavioral data that can validate our norms. But in lieu of such data, perhaps the first step in data validation should be to examine the well-studied effect of frequency on reduction at the word and segment level (Aylett and Turk 2006; Gahl 2008). Indeed, we are currently working on a spoken version of the corpus used for this study, and so appropriate data will soon be available for probing the effects of frequency on reduction patterns. We hope that with this, and other efforts, future research will expand the empirical base for investigating frequency in Amazigh languages.

Appendix

All of the frequency data and scripts for generating them are available at the first author's website. Two data tables are especially useful. The file <wordfreq_master> (in various formats) gives all of the information relevant to word frequency and word structure documented in section "Word frequency". This file is the lexicon, in effect, listing all words as rows. The seven columns in the document analyze words for the following attributes: lexical representation, phonetic representations (including all variant forms), morphological gloss, part of speech category, the number of morphemes contained in the word, word frequency, and word probability.

The file <soundfreq_master> contains all of the information for investigating sound frequencies discussed in section "Consonant and vowel frequency". It has 75 rows (for 36 singleton consonants, 36 geminates, and three vowels) and 20 columns that give the frequencies for these 75 segments cross-classified by all of the categories discussed here (i.e., level of representation, frequency type, morpho-syntactic categories). Many levels of structure can be investigated: for example, token frequency in nouns, or open class words (both nouns and verbs), or all words can be easily extracted.

The Python scripts for investigating the corpus are also available from the project page. The Python notebooks for these scripts are fully commented and enable users to replicate the results reported here—and extend them to new projects.

References

- Aksan Y, Yaldir Y (2012) A corpus-based word frequency list of Turkish: evidence from the subcorpora of Turkish National Corpus project. *Studia Uralo-altaica* 49:47–57
- Anand P, Chung S, Wagers M (2011) Widening the net: challenges for gathering linguistic data in the digital age. White Paper for the National Science Foundation series NSF SBE 2020: future research in the social, behavioral, & economic sciences (White Paper ID 121)
- Aylett M, Turk A (2006) Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *J Acoust Soc Am* 119:3048–3058
- Baayen RH, Piepenbrock R, Gulikers L (1996) CELEX2. Linguistic Data Consortium, Philadelphia
- Beckman JN (1999) Positional faithfulness: an optimality theoretic treatment of phonological asymmetries. Garland, New York

- Bensoukas K (2014) Feature dissimilation in Tashlhit. University of Mohammed V, Publications of the Faculty of Letters and Social Sciences, Rabat [These and Memoirs series, no. 72]
- Blevins J (2004) Evolutionary phonology: the emergence of sound patterns. Cambridge University Press, Cambridge
- Bod R, Hay J, Jannedy S (2003) Probabilistic linguistics. The MIT Press, Cambridge, MA
- Boukous A (1977) Langage et culture populaire au Maroc, essai de sociolinguistique. Dar El-kitab, Casablanca
- Boukous A (1987) Phonotactique et domaines prosodiques en berbère (parler tachelhit d'Agadir, Maroc) (Doctoral dissertation). Université Paris 8
- Boukous A (2012) Revitalisation de la langue Amazighe. Defis, enjeux et strategies. Top Press, Rabat
- Bybee J (2001) Phonology and language use. Cambridge University Press, Cambridge
- Connine CM, Ranbom LJ, Patterson DJ (2008) Processing variant forms in spoken word recognition: the role of variant frequency. *Percept Psychophys* 70:403–411
- Dell F, Elmedlaoui M (2002) Syllables in Tashlhiyt Berber and in Moroccan Arabic. Kluwer, Dordrecht
- Gahl S (2008) Time and thyme are not homophones: the effect of lemma frequency on word durations in spontaneous speech. *Language* 84:474–496
- Gerz D, Vulić I, Ponti E, Naradowsky J, Reichart R, Korhonen A (2018) Language modeling for morphologically rich languages: character-aware modeling for word-level prediction. *Trans Assoc Comput Linguist* 6:451–465
- Gries ST (2016) Quantitative corpus linguistics with R: a practical introduction. Routledge, New York
- Hayes B, Wilson C (2008) A maximum entropy model of phonotactics and phonotactic learning. *Linguist Inq* 39:379–440
- Jebbour A (1996) Contraintes prosodiques et morphologie en berbère (tachelhit de Tiznit – Maroc). Analyse linguistique et traitement automatique (These de Doctorat d'Etat). Université Mohammed V, Rabat, Morocco
- Jebbour A, Boukous A, El Alami A, Li JS, Ridouane R, Alderete J (2021) Nine texts of Tashlhiyt Berber: structured representations of 18,000 words (first release)
- Kessler B, Treiman R (1997) Syllable structure and the distribution of phonemes in English syllables. *J Mem Lang* 37:295–311
- Levelt WJM, Wheeldon L (1994) Do speakers have access to a mental syllabary? *Cognition* 50: 239–269
- Orzechowska P, Ridouane R (2018) The structure of vowelless verbal roots in Tashlhiyt Berber. Paper presented at the 9th international conference on speech prosody 2018, Poznan, Poland
- Pitt MA, Dilley L, Tat M (2011) Exploring the role of exposure frequency in recognizing pronunciation variants. *J Phon* 39:304–311
- Whaley CP (1978) Word-nonword classification time. *J Verbal Learn Verbal Behav* 17:143–154
- Willerman R (1994) The phonetics of pronouns: articulatory bases of markedness (Doctoral dissertation). University of Texas at Austin
- Zipf GK (1949) Human behavior and the principle of least effort. Addison-Wesley, Reading



On State and Case in Berber: A Typological Perspective 22

Aicha Belkadi

Contents

Introduction	457
The Morphosyntax of the Alternation: Diachrony, Synchrony, and Micro-variations	459
Morphology	459
Common Syntactic Patterns of the States	461
Syntactic Micro-variations	464
Accounts of the Berber Nominal Alternation in the Light of Typological Data	466
Construct State	467
Configurational Coding	469
A New Unrecognized Category	472
Marked Nominative	474
Challenges to Case in Berber and Typology	476
Post-Verbal S and A with Quantifiers	477
O Arguments in the Annexed/Nominative Form	479
Adjectives and Lack of Case Agreement	483
Conclusion	487
References	488

The materials in this paper have been presented to several audiences. Their comments and suggestions are hereby gratefully acknowledged. In particular, I wish to thank the members of the Surrey Morphology Group, who directed me to languages, morphological constructions, and sources in the literature that greatly helped this work. I am also very grateful to Dr. Lameen Souag, who read a draft of this paper and was generous enough to share his rich knowledge with me. Finally, I also thank my Taqbaylit speaking colleague Dr. Samir Ben Si Said, my consultant Arezki Salem, and all my informants who have helped with the Taqbaylit data.

A. Belkadi (✉)
SOAS University of London, London, UK
e-mail: ab105@soas.ac.uk

Abstract

The Berber system of nominal marking – known as the annexed vs. free state alternation – is one of the most researched aspects of the phylum's morphosyntax. While the surface properties of the two states are well-understood, there still exists no consensus on their precise function(s). In this paper, I consider the issue from a typological perspective and show that the Berber state alternation is better analysed as a marked nominative case marking (following König and Aikhenvald). I focus specifically on three arguments that are often used in Berber scholarship to dismiss the case analysis and defend other types of hypotheses: (i) the peculiar behaviour of nominal states with some quantifiers; (ii) the fact that some direct objects can be marked in the same way as subjects; and (iii) the lack of case agreement on attributive adjectives. Based on cross-linguistic data and theories, I argue that these arguments do not actually challenge a marked nominative case analysis, but rather support it.

Abbreviations

A	Agent
ABS	Absolutive
ACC	Accusative
ADE	Adessive
AGEN	Agent
ANAPH	Anaphor
ANP	Anaphoric Deictic Clitic
AOR	Aorist
AS	Annexed State
CAUS	Causative
COP	Copula
CS	Construct State
CSTR	Construct
DAT	Dative
DEF	Definite
DEM	Demonstrative
EA	Etat Annexé
ERG	Ergative
EVID	Evidential
F	Feminine
FS	Free State
GEN	Genitive
INTJ	Interjection
IRR	Irrealis
LOC	Locative
M/Ma	Masculine
NEG	Negation

NEUT	Neutral
NSG	Non Singular
OBJ	Object
P	Patient
PAR	Partitive
PFV	Perfective
PL	Plural
POSS	Possessive
POT	Potential
PRF	Perfect
PRO	Pronoun
PROX	Proximal
PRS	Present
PRT	Particle
PTCP	Participle
S/SG	Singular
VEN	Ventive

The system of nominal marking, known as the state alternation, is, undoubtedly, one of the most researched aspects of Berber morphosyntax. Its surface properties are well-understood: depending on a range of factors, including grammatical functions and word order, nouns occur in one of two morphological forms, referred to as the free state and the “annexed” or “construct” state. In (1a), for instance, the noun *ufunṭum* “ghost” surfaces in the annexed state because it is the clause subject and follows the verb. In (1b), by contrast, it occurs in pre-verbal position and as such occurs in the free state form *afunṭum* (“ghost”).

a. t-ənnə=yas y-əffər u-funṭum dəffir n təjra
3SGF-tell.PFV=3SGM.DAT 3SGM-hide.PFV AS-ghost behind of tree
“She told him that a ghost was hiding behind the tree.”

b. a t-əttxayal a-funṭum nni y-əffər dəffir n təjra
PRT 3SGF-imagine.IMPF FS-ghost DEM 3SGM-hide.PFV behind of tree
“She is imagining that this ghost is hiding behind the tree.”

Analyses on the function of these nominal states are diverse. In some parts of the literature, the states are analyzed as case markings (Aikhenvald 1995; Ouhalla 1993; Ennaji 2001; König 2006; Mettouchi 2008). In other parts of scholarship, the marking is considered to be a unique morphosyntactic system of coding, designating a relation between a noun and some other elements preceding it, including nouns and agreement morphemes, or clitics (Galand 1969; Chaker 1988; Mettouchi and

Frajzyngier 2013). Finally, in the analyses developed within the transformational framework, the states are argued to represent the formal syntactic configuration in which a noun occurs, and, or the different levels of functional structure it projects (Guerssel 1992, 1995; El Moujahid 1997; Bendjaballah and Haiden 2005; Achab 2006; El Hankari 2014).

While no real consensus exists, all these analyses agree that the annexed state is linked, in some way or another, to dependency. Dependency marking is very much what case is. But if, superficially, the alternation has all the hallmarks of case, there are many ways in which it contrasts from prototypical views of what case is and how it should behave. These “non-canonical” properties raise the question whether the nominal alternation is really case or if, as proposed by some scholars, such as Mettouchi and Frajzyngier (2013), it constitutes a type of unrecognized morphosyntactic phenomenon. From a typological perspective, categorization of this phenomenon has important ramifications. Whether it is case, a configurational type of marking or a new morphosyntactic category, its non-canonical properties, some of them quite rare, need to be acknowledged, explained, and shared with our scientific community.

My principal objective in this chapter is to contribute to the already rich scholarship on the issue, adopting a typological perspective. I consider research on similar phenomena across the languages of the world and the extent to which these can enhance our understanding of the Berber states. In particular, I defend the marked nominative analysis of the state alternation, which contends that the free state and the annexed state represent, respectively, accusative case and marked nominative case. I focus on three arguments that are often used in the literature to dismiss case analyses and defend other types of hypotheses: (i) the peculiar behavior of nominal states with some quantifiers; (ii) direct objects marked in the same way as subjects; and (iii) the lack of case agreement on attributive adjectives. Based on cross-linguistic theories and data, I show that these arguments do not actually challenge a marked nominative case analysis, but rather support it. My second goal here is to provide a pan-Berber description of the state alternation, including discussion of relevant micro-variations in the syntactic distribution of the states, as well as an evaluation of the different types of morphosyntactic analyses of the alternation, again, in light of typological data. This description, which is certainly not the first, is aimed principally at typologists and descriptive linguists who are not familiar with Berber grammar, literature, and terminology, as well as students new to the field of Berber linguistics.

The chapter is organized as follows: The section titled “[The Morphosyntax of the Alternation: Diachrony, Synchrony, and Micro-variations](#)” describes the morphosyntax of the state alternation in the Berber languages in which it is found, including a discussion of micro-variations. Section “[Accounts of the Berber Nominal Alternation in the Light of Typological Data](#)” critically evaluates the four main types of analyses of the state alternation that have been proposed in the literature. And finally, section “[Challenges to Case in Berber and Typology](#)” focuses on the marked nominative approach and addresses three challenges it faces.

The Morphosyntax of the Alternation: Diachrony, Synchrony, and Micro-variations

In diachrony, the quasi-universality of the state alternation across Berber languages is undoubtable. In synchrony, however, the nominal opposition remains productive only in Tuareg, all varieties of Morocco, including Tarifit, Tashelhit, and Tamazight, and most of the languages of Algeria, including Chenoua, Taqbaylit, Tashawit, Tasaḥlit, Mzabi, and Gourara (Kossmann 2002; K'halainou 2006; Mettouchi 2014; Garaoun 2021). Other Berber languages, such as those traditionally classified in the Eastern Berber group, for example, Siwi (Egypt), Ghadamsi, Awijilah (Libya), and Djerba (Tunisia), as well as several varieties from the Southern group, i.e., Zenaga (Mauritania), only retain morphological traces of the alternation, mainly in topographic nouns or old folk songs (Lanfry 1971; Brugnatelli 1987; Souag 2010; Kossmann 2013). (In Ghadamsi, some nouns appear in the annexed state after prepositions in certain folk songs (Lanfry 1971–1972: 181; Kossmann 2013: 20). In Zenaga and Siwi, traces of the state opposition are found in topographic names. Brugnatelli (1987: 349), for instance, cites the following nouns of places: “Qurat Uazidi” in Siwa, in which *uazidi* is the annexed form of *azidi* “wolf,” and *ə-n-wass* in Zenaga, where *wass* is the annexed form of *ašš* “day.”)

The morphophonological and morphosyntactic properties of the alternation have been extensively described for individual languages (Chaker 1983; Ouhalla 1996; Kossmann 1997; Achab 2003; Bendjaballah and Haiden 2005; Mettouchi and Frajzyngier 2013; Lahrouchi 2013; El Hankari 2014; Ben Si Said 2020, among others), but also from pan-Berber perspectives (Chaker 1988; Kossmann 2000; Mettouchi 2014). These studies have shown great commonalities in the distribution and characteristics of the annexed and free states across Berber, as well as some micro-variations; particularly the study by Mettouchi (2014), which offers the premises for a typology of the states based on four Berber languages. The aim of this section is two-fold: (i) to present the morphosyntactic properties of the states across Berber, with a special focus on synchrony, diachrony, and micro-variations, and (ii) make the morphosyntax of the Berber state alternation more accessible to cross-linguistic typological research.

Morphology

The state of a noun in Berber is mainly marked by vocalic and semi-vocalic prefixes, whose forms depend on the gender and number of a noun. The free state masculine singular is largely considered to be the default form of Berber nouns, based on which the annexed state and the plural free state are derived (Basset 1952; Galand 1964; Chaker 1998; Guerssel 1992; Bendjaballah and Haiden 2005; Achab 2006; Lahrouchi 2011; Ben Si Said 2014). In the Northern branch languages, the free state masculine singular is marked by the prefixes **a-**, **u-**, or **i-**. In Tuareg varieties, which have a larger vowel inventory, other prefixes are

found. The plural free state is most commonly marked by the prefix **i-**, but many nouns, especially those whose free state singular is marked by the prefix **u-**, keep the same vocalic prefix in the plural.

The formation of the annexed state is slightly more complex. For masculine nouns, it depends on the morphophonological structure of a noun, the language, and often even the particular variety of a language. In most varieties of the Northern group, it is marked either by (i) a change of the initial free state vowel – **a-** becomes **u-**; (ii) the deletion of the initial free state vowel and addition of a semi-vowel – **w-** in the singular and **y-** in the plural –; or (iii) the addition of a semi-vowel without deletion of the initial free state vowel (Brugnatelli 1987; Kossmann 1997; Lahrouchi 2013; Bendjaballah and Haiden 2005; Ben Si Said 2014; Mourigh 2015, among others). (In several Taqbaylit varieties, the glides **w-** and **y-** can be realized, respectively, as the segments /p/ or /b/ and /g/. According to speakers, the segment /p/ is associated with women, /b/ with men. These allophones are found in contexts in which a noun in the annexed state is preceded by the genitive preposition *n* (cf. section “Common Syntactic Patterns of the States”).) There are, of course, variations in the formation of the annexed state. Thus, there are nouns whose annexed form is marked by a drop of the initial vowel or whose free state initial vowel, especially if it is **u-**, remains unchanged. In Touareg, while some remnants of the semi-vowels can be found, the annexed state formation principally involves the deletion, or phonological reduction, of the initial vowel (Prasse 1974; Brugnatelli 1987; Heath 2006; Kossmann 2013). Some examples of derivations are provided in (2) and (3), below. Taqbaylit represents the formation in most Northern varieties, and Ayer Tuareg represents the formation in Touareg languages.

(2) Formation of the masculine annexed state in Taqbaylit (Algeria; author’s own corpus)

	Free State	Annexed State	
Singular	aqəcwal əyyul aman	uqəcwal wəyyul waman	‘basket’ ‘donkey’ ‘water’
Plural	irgazən	yərgazən	‘men’

(3) Formation of the masculine annexed state in Ayer Tuareg (Niger; Kossmann 2013)

	Free State	Annexed State	
Singular	əyālak əyiwān edāg	āyālak əyiwān ādāg	‘people’ ‘camp’ ‘place’
Plural	iγərman	γərman	‘cities’

The free state of feminine nouns follows the same patterns as that of masculine nouns, in both the singular and plural, with the addition of the feminine circumfix marker. The feminine annexed state is generally marked by a drop of the initial

vowel. The following examples from Taqbaylit show how the annexed state of feminine nouns is formed.

(4) Formation of the feminine annexed state in Taqbaylit (Algeria; author’s own corpus)

Free State	Annexed State	
taqəcwalt	tqəcwalt	‘little basket’
taqciɛt	təqciɛt	‘girl’
tifirəst	tfirəst	‘pears’

From a diachronic perspective, the free state and annexed state prefixes are generally agreed to have grammaticalized from a set of old demonstratives (Stumme 1899; Vycichl 1957; El Moujahid 1993; Brugnatelli 1987; Achab 2006; Ben Si Said 2014). Vycichl (1957) and Brugnatelli (1987) note the formal similarities between the most commonly found prefixes, **w-** and **a-**, and the demonstratives *wa*, *ta*, *wi*, *ti* – found across various Berber languages. They hypothesize that, at a certain point, these demonstratives developed into articles, which at some point attached to the noun and lost their article function. A slightly divergent analysis on the diachrony of the state markers is presented by Prasse (1974: 14), who argues that the annexed state semi-vowel is developed from a former pronoun.

Common Syntactic Patterns of the States

Some languages use the annexed state only in a subset of domains. It is the case of Ghomarra (Morocco), for example, in which the annexed state only occurs in the prepositional domain, and in noun phrases headed by the numeral *yan/yat* ‘one’ (Mourigh 2015). In most other languages in which it is productive, the contrast between the annexed state and the free state is meaningful in the prepositional, nominal, verbal, and clausal domains.

A noun occurs in the annexed form if it is (i) the post-verbal subject of an intransitive (S) or transitive verb (A), (ii) the argument of a preposition, and (iii) [in some languages] as the dependent of a nominal head. These distributions are illustrated by the following examples.

(5) Post-verbal intransitive subject (S)
Tamasheq (Mali, Heath 2006: 574)¹

ì-wæt ǎ-jənnə
3MaSgS-hit.PerfP Sg-rain
“The rain struck (=fell).”

¹(All examples that are taken from other sources in the literature are transcribed and glossed as per the source.)

(6) Post-verbal transitive subject (A)

Taqbaylit (Algeria, author's own corpus)

y-ufa=d w-əqɛic nni tabəllart nni inəs d tiləmt
 3SGM-find.PFV=VEN AS-boy this jar.FS this POSS.3SG COP empty
 “The boy found his jar empty.”

(7) Dependents of prepositions

a. Ghomarra (Morocco, Mourigh 2015: 63)

i-ḡḡ=aheṇ ḡ u-qemmum nn-es
 3sgm-do:P=S:ANP in ms:a-mouth of-3ms
 “He put them in his mouth.”

b. Tashawit (Algeria, Penchoen 1973: 34)

t.ṭṭ.f i-ḡḡ.n̄s ufus afusi
 “She holds the other with her right hand.”

Inside noun phrases, dependents generally follow the genitive preposition *n* “of,” and as a result also occur in the annexed state. This includes possessors (8), nominal complements (9), as well as dependents of several quantifiers, including numerals (10).

(8) Tashawit (Algeria, Penchoen 1973: 19)

umas n-tslit am.zzyan
 brother of-bride.as small
 “The bride’s little brother.”

(9) Ghomarra (Morocco, Mourigh 2015)

lqaḍiya n txwraft
 case of story.EA
 “The case of the story.”

(10) Taqbaylit (Algeria, author's own corpus)

arrac nni, ṭlata nni n̄ w-arrac, wwi-n yiwət n̄ ṭfirəst,
 children DEM, three DEM of AS-children, take.PFV-3PLM one of pear
 “Those children, those three children, took a pear each.”

Some varieties seem to omit, in certain contexts at least, the genitive preposition. In these noun phrases, as shown in the following examples, dependents remain in the annexed state.

a. aglim *our'ilas*
skin panther
"The skin of the panther."
b. thifounasin *oufella*
Cows farmer
"The farmer's cows."

axxam umzzian
house small
“The house of the small one.”

y-əseə *a-mqarqur* daxəl n tbəllaɾt
3SGM-have.PFV FS-frog inside of jar
“He had a frog inside a jar.”

təksi dəgʏa ddqər (,) *tasrit-nnəs*
 “She soon became pregnant, her daughter-in-law.”

d	<i>a-mcic</i>
cop	fs-cat
“It is a cat.”	

(16) Prepositions and quantifiers

a. **Tamasheq (Mali; Heath 2006)**

əɾjsæš-æɣ =d hār *a-mézzaɣ*.
 walk.PFV-1SG =VEN until FS-camp
 “I walked until the camp.”

b. **Taqbaylit**

kul *a-rgaz*
 all FS-man
 “Every man.”

Syntactic Micro-variations

There are, across Berber, interesting micro-variations in the marking of subjects, particularly S, and direct objects. As will be discussed in more detail in subsequent sections, these micro-variations, which affect core arguments, play an important role in the various analyses presented to explain the nominal alternation, and, particularly, they play a crucial role in whether the alternation can be described as a case or not.

The first type of constructions in which variations in the marking of nouns can be observed are copula clauses: in clauses whose main predicate is a noun or adjective preceded by the copula *d*, S arguments can be overtly realized. In such cases, S usually precedes the *d* + predicate complex and occurs in the free state form, as shown by the Figuig example in (17).

(17) **Figuig (Morocco; Kossmann 1997)**

a-qšiš a, d əmmi-ç.
 FS-child this COP FS.son-POSS:2SGM
 “This child, it is your son.”

Taqbaylit, Tamazight, and Tarifit, however, allow S arguments to follow the predicate (Mettouchi 2011; Mettouchi and Frajzyngier 2013; Taifi 2012; Lakfioui 2000). In Taqbaylit and Tarifit, a non-initial S can be in the annexed state, as shown by examples (18a) and (18b). In Tamazight, on the other hand, S is always in the free state. This is illustrated in (19).

(18) a. **Taqbaylit (Mettouchi and Frajzyngier 2013)**

d t-a-gugam-t *t-q/iɣ-t* nni
 COP FEM-FS-mute-FEM FEM-AS.child-FEM DEM
 “That girl is/was mute.”

b. **Tarifit (Morocco; Lafkioui 2000)**

qad tamza(.) *tnni*
 COP ogress AS.PRO.DEM.FEM
 “She is an ogress, this one.”

(19) **Tamazight (Morocco; Taifi 2012: 259)**

- a. *aserdun-a*, d aḥrun.
 FS-mule-DEM COP unruly
 “This mule is unruly.”
- b. d aḥrun, *a-serdun-a*.
 COP unruly FS-mule-this
 “This mule is unruly.”

A very similar type of constructions are right-dislocations. Across Berber, right-dislocated S and A are usually in the free state. In Taqbaylit, right-dislocated S and A can be in the annexed state (as discussed in section “[Common Syntactic Patterns of the States](#)”) but can also surface in the free state.

(20) **Taqbaylit (Algeria; author’s own corpus)**

- wala-γ=tt, *taməṭṭut*.
 see.PFV-1SG=3SGF:ACC woman.FS
 “I saw her, the woman.”

“Pseudo-verbal” constructions (after Kossmann 1997) are other contexts in which important variations occur. A pseudo-verbal construction is one in which a predicate shares some properties of verbs, but not others. Precisely, these predicates take some form of agreement but are not inflected for aspect. The agreement marker on the predicate is usually an accusative clitic. In most varieties, S is in the free state. In Taqbaylit, though, in some pseudo-verbal constructions, S can be either in the annexed state or the free state depending on its position. It is in the annexed state if it follows the predicate and in the free state if it precedes it.

(21) **Taqbaylit (Mettouchi and Frajzyngier 2013)**

- diri=t *wə-rgaz* inna
 bad = ACC:3SGM AS-man DEM.DIST
 “That man is bad.”

(22) **Taqbaylit (Basset and Picard 1948)**

- a-dil* agi, dir=it i wuči
 FS-grape DEM bad=ACC:3SGM to eating
 “This grape is not good for eating.”

Somewhat similar to the construction above are clitic-doubled object constructions. In most Taqbaylit varieties, and some Tamazight varieties (Guerssel 1995), a direct object can be doubled by an accusative clitic. Clitic-doubled direct objects, which follow the verb, surface in the annexed state.

(23) **Taqbaylit (Bendjaballah and Haiden 2005)**

- a. jə-ttʃa *a-çsum* nni
 3SGM-eat.PFV FS-meat DEM
 “He ate that meat.”
- b. jə-ttʃa $\overline{=t}$ *wə-çsum* nni
 3SGM-eat.PFV =ACC:3SGM AS-meat DEM
 “He ate that meat.”

The clitic-doubled direct object in annexed state may be inside the core clause (Chaker 1983, 1998), as above, or occur after a prosodic boundary (Mettouchi 2011), as below.

(24) **Taqbaylit (Mettouchi 2011)**

- t-ufa d a-mʃiʃ n wə-drar //
 3SGF-find.PFV COP FS-cat of AS-mountain //
 “She found it was the mountain cat”
- i= \overline{t} i-zədy-ən / *wə-xxam* nni.
 COMP=ACC.3SGM 3SGM- inhabit.PFV-PTCP / AS-house DEM
 “who inhabited it, the house.”

Accounts of the Berber Nominal Alternation in the Light of Typological Data

It is evident from the descriptions in the previous sections that the state alternation correlates with the phenomenon of dependency. The annexed state is exclusively found on nominals which are functionally dependent on a head while nouns that are not functionally dependent on other heads systematically surface in the free state form. All analyses of the state alternation, therefore, center on dependency in one way or another. Of course, they explain the link between the Berber nominal states and dependency in different ways: either configurationally or functionally, as a type of case, or as another type of morphosyntactic coding. Many of these accounts rely on analyses of similar phenomena in other languages, or the tools and architectures of particular syntactic frameworks. While describing all previous proposals on the subject is outside the scope of the present chapter – and has been done in previous work on the topic (cf., for instance, Mettouchi and Frajzyngier 2013; or El Hankari 2014) – providing an overview of some accounts is important and highly relevant to the discussion in section “Challenges to Case in Berber and Typology.” On the one hand, that analyses are so often inspired by research on various morphosyntactic phenomena in other languages shows the versatility of the Berber state alternation. On the other hand, that it continues to be a morphosyntactic puzzle, subject to debate, shows how the Berber state coding is unique and does not canonically fit with any known typological phenomenon.

Construct State

The “state” terminology widely used to describe the Berber system of nominal coding under discussion is borrowed from the Semitic linguistics literature. In Semitic languages, nouns famously receive a special morphological marking, referred to as the construct state or construct form, when they occur in some noun phrases, often described as bare genitive noun phrases (Ritter 1991; Fassi-Fehri 1999; Engelhardt 2000; Benmamoun 2000, among others). Consider the following Hebrew examples.

(25) Hebrew (Semitic; Ritter 1991)

- a. **ha-bayit** fel ha-mora
DEF-house of DEF-teacher
“The house of the teacher.”
- b. **beyt** ha-mora
house.CS the-teacher
“The teacher’s house.”

In (25a), the nominal head *ha-bayit* “house” is linked to its dependent noun *ha-mora* “teacher” by the preposition *fel* “of”. In (25b), by contrast, the preposition is omitted and the head noun occurs in another form, *beyt*. The special coding of the noun in (25b) is referred to as the construct form, while the form in (25a) is generally referred to as the free state, or free form.

The adoption of the Semitic terminology reflects an assumption and implication that the Berber state alternation is an instance of construct form à la Semitic. (Although, as rightly pointed out by one of the reviewers, Basset (1932) states that the terminology “free” vs “annexed” (“libre” vs “annexé” in French) is not very appropriate but chosen for lack of a better terminology.) The Berber and Semitic annexed/construct forms do indeed share some obvious semantic and syntactic similarities. Semantically, they may entail a relationship defined by possession while syntactically they involve the same word order, in which the head noun precedes its dependent. Compare, for instance, the following Berber example with the abovementioned Hebrew example.

(26) Taqbaylit (Author’s own corpus)

- a-xxam n w-ərgaz
FS-house of AS-man
“The man’s house.”

More recently, some analyses have also relied on this parallel to explain the Berber state alternation. Ouhalla (1996) and Ennaji (2001), for instance, have adopted proposals in the Semitic research, developed within transformational and minimalist frameworks, which analyze the construct state as a marker of genitive case assigned in specific structural configurations. Ouhalla (1996), in particular, proposes that this genitive case is assigned in Berber in the specifier positions of an agreement phrase, AgrP.

An analysis of the Berber annexed state, as a type of construct state, appears genetically and typologically coherent. The construct state is not just a Semitic construction but constitutes, in fact, a very frequent cross-linguistic phenomenon, found in various languages in Africa, America, Eurasia, and Oceania (Creissels 2009, 2017; Rießler 2016). And many of the African languages with construct state codings are from the Chadic and Cushitic genera which, like Berber and Semitic, belong to the Afro-Asiatic phylum. In Hausa, the construct state marking is even more similar to the Berber annexed state, since the affix also encodes gender and number, as shown in the following examples.

(27) **Hausa (Chadic; Afro-Asiatic; Creissels 2017: 6)**

- | | | | |
|----|----------------|--------------|-------|
| a. | kàre-n | | Daudà |
| | dog-CSTR.SG.M | | Dauda |
| | “Dauda’s dog.” | | |
| b. | kàree | na | Daudà |
| | dog | that_of.SG.M | Dauda |
| | “Dauda’s dog.” | | |
| c. | saaniyar | | Daudà |
| | cow-CSTR.SG.F | | Dauda |
| | “Dauda’s cow.” | | |
| d. | saaniyaa | ta | Daudà |
| | cow | that_of.SG.F | Dada |
| | “Dauda’s cow.” | | |

The Berber annexed state, however, differs from the construct state in a number of crucial ways. First, as explained by Creissels (2009, 2017), in all the languages in which it is found, the construct involves a morphophonological modification of the nominal head, not its dependent. This fact can be observed in the previous Hebrew and Hausa examples in which it is the head of the noun phrases, respectively, *beyt* “house,” *kàre-n* “dog,” and *saaniyar* “cow,” that occurs in the construct form. In Berber, the annexed state is marked on dependent nouns, such as *wərgaz* “man” in example (27) above.

Second, the construct form is always found inside noun phrases. This can be clearly observed in the following example from Persian. The noun *xâne* “house” is in the construct form when it heads a noun phrase but in the free form when it is the subject of the clause.

(28) **Persian (Mahootian 1997, cited in Rießler 2016: 18)**

- | | | | |
|----|---------------------|-----------------|------------|
| a. | xâne-ye | | bozorg |
| | house-construct | | big |
| | “Big house.” | | |
| b. | in | xâne | bozorg ast |
| | dcm | house(absolute) | big is |
| | “The house is big.” | | |

The Berber annexed state, by contrast, occurs in a much wider range of syntactic contexts—after prepositions, and, crucially, in post-verbal subject positions (cf. sections “[Common Syntactic Patterns of the States](#)” and “[Syntactic Micro-variations](#)”). Finally, in Semitic and other languages, the annexed state is in complementary distribution with prepositions, case markers, or other functional markers. In Berber, on the other hand, nouns are in their annexed forms, whether or not the genitive preposition *n* is omitted.

Incidentally, many languages have so-called anti-construct state constructions, in which, as in Berber, it is the dependent inside a noun phrase that is marked with the construct. This is illustrated below in example (29a) from Kildin Saami, in which the adjective *ēllis* “high” is in the construct instead of the head.

(29) **Kildin Saami** ([Uralic Rießler 2016: 42](#))

- | | | | | |
|----|-------------------------|--------|-----------|--------|
| a. | Tedt | lī | ēl'l'-es' | pērrht |
| | DCM | COP | high-ATTR | house |
| | “This is a high house.” | | | |
| b. | Tedt | pērrht | lī | ēll |
| | DCM | house | COP | high |
| | “This house is high.” | | | |

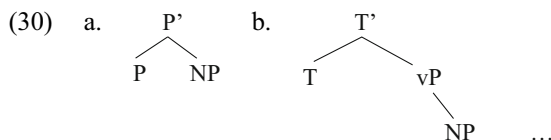
But again anti-constructs differ from Berber state distinction. Like the construct, the anti-construct is found exclusively inside noun phrase constructions. Furthermore, in anti-constructs, as shown in (29) attributive adjectives are marked, like all noun phrase dependents. In Berber, adjectives and non-nominal dependents are not marked.

Configurational Coding

Among the most frequent analyses of the alternation in modern Berber scholarship, configurational coding are those which link the two nominal states to different syntactic configurations in which a noun can occur. All of these configurational accounts are framed within some version of Chomsky’s Minimalist framework and adopt the premises that the surface properties of a sentence and functional relations between different elements in that clause, such as grammatical relations, are mediated by hierarchical tree structures.

Many of these studies analyze the annexed state as the form taken by a noun being “c-commanded” by another head (El Moujahid [1997](#); Bendjaballah and Haiden [2005](#); El Hankari [2014](#)). These studies actually formalize a specific characteristic of the annexed state, on which previous descriptive accounts, such as those of Basset’s ([1952](#)) and Applegate’s ([1970](#)) focus: the fact that it occurs on nouns which follow specific heads such as verbs, prepositions, and in some Berber varieties, nouns. El Hankari ([2014](#)), for example, proposes that the annexed form is given to nouns that are c-commanded by a prepositional head P, or the tense head T. El Hankari’s configurations are illustrated in (30) below. According to fig. (30a),

nouns which follow prepositions are in the annexed state. Under the configuration in (30b) any post-verbal subject, assumed to be in the specifier position of the vP constituent, which is c-commanded by T occurs in the annexed state.



Other types of configurational accounts analyze the distinction between states as differences in the internal syntactic structure of the phrases headed by nouns (Guerssel 1992, 1995; Achab 2006; Bendjaballah and Haiden 2005). Bendjaballah and Haiden (2005), focusing particularly on the relation between prepositions and their complement nouns, propose that the annexed state reflects a noun's phonological and syntactic deficiency, as well as its phonological and syntactic dependence on a preceding c-commanding prepositional head. Put simply, a noun in the annexed state projects less structure, and as such needs to be structurally combined with another head. Guerssel (1992, 1995) proposes an analysis along similar lines. He attributes the state alternation to the different configurations in which case can be marked. Particularly, he proposes that Berber nominal projections, which he assumes to be DPs, occur within a larger functional projection, which he terms KP, headed by case markers and evidently associated with the assignment of case (cf. 31).

(31) $[_{KP}K[_{DP}]]$

K^0 , the head of KP can be occupied by a range of case markers which include all prepositions preceding a noun in annexed form and several default markers, or it can be left empty. The prepositions occurring in K^0 are each markers of a specific case, such as dative, genitive, comitative, etc., while the empty KP case-marks post-verbal noun phrases. The default markers occurring in K^0 mark accusative case as well as default case (e.g., when the subject is left-dislocated), and correspond to the first vowel of a Free State noun. Prepositions associated with FS nouns are, in turn, taken to occur as real prepositions in a higher PP. The range of examples in (32) below illustrate, with the noun *zru* "stone," the various ways in which case can be marked within this analysis (Guerssel 1992: 117).

- (32) a. $[_{KP} \quad s \ [wzru]]$ 'with a stone' (Case marker: instrumental *s*)
 b. $[_{KP} \quad a \ [zru]]$ 'stone' (Case marker: default *a*)
 c. $[_{KP} \quad \emptyset[wzru]]$ 'stone' (Case marker: empty)

An analogous proposal is put forward by Achab (2003, 2006). Nouns which occur within an NP projection are argued to occur in their annexed form while those projecting an additional portion of structure occur in the free state form. In his account, the free state vowel is a determiner, heading a determiner phrase (DP).

According to this proposal, the two forms of the noun “man” *argaz*, and *wərgaz*, can be represented as follows:

- (33) a. Free State b. Annexed State
- DP

├── D

└── NP

|

N

a rgaz

NP

└── N

|

wrgaz

On the surface, however, annexed state nouns do not occur as “deficient” NPs, but as complex DPs headed by either a preposition or, when they are post-verbal subjects, the agreement marker affixed onto the verb. This is illustrated in the following tree structures:

- (34) a.
- VP

├── V

└── DP

| └── D └── NP

| | |

agri ti N

|

wrgaz
- b.
- DP

├── D

└── NP

| |

prep N

|

wrgaz

Similarly to Guerssel (1992, 1995), Achab (2006) positions prepositions which do not occur with the annexed state in an independent PP projection, selecting DP as their complement.

- (35)
- PP

├── P

└── DP

| └── D └── NP

| |

prep a- N

|

rgaz

All these proposals provide formalizations and explanations of the function of the states that are, overall, strongly dependent on the assumptions made by the Minimalist framework about the structure of human language. Indeed, most of the

concepts and configurations that account for the distribution of the states here are very specific to this model. The operation of “c-command” or the idea that case should be configurationally-assigned, for instance, adopted to explain why dependents of prepositions or post-verbal subjects occur in the annexed state, have no explanatory values outside of Minimalism.

Furthermore, in both Guerssel and Achab’s accounts, there is no evidence that different prepositions occupy different structural slots except from the particular marking on a noun. If a preposition does not govern the annexed state, then it is assumed not to be in D or in K, but in P. The only reason to do so is that the free state vowel is assumed to be in D or K. Interestingly, c-command relations, which seem central to many of these configurational accounts, are, among other things, associated with the assignment of case in Minimalism. I will come back to case in section “[Marked Nominative](#).” Beforehand, I describe another proposal of the state alternation.

A New Unrecognized Category

Mettouchi and Frajzyngier (2013) develop an account of the state alternation, based on Taqbaylit (Algeria). They propose that the annexed state occurs on nouns which provide a value for a grammaticalized function carried by a preceding element (Mettouchi and Frajzyngier 2013). This proposal is slightly reminiscent of a previous analysis by Basset (1932), Galand (1964), and Chaker (1988), who argue that the annexed state marks a noun which completes the interpretation of, or is dependent on, a preceding nominal such as a preposition, a noun, or a person marker on the verb. Mettouchi and Frajzyngier’s (2013) account is relevant, not just to work on Berber languages, but also to linguistic typological research. Indeed, the state alternation they describe is claimed to be a unique, previously undescribed, morphosyntactic phenomenon. According to their analysis, certain morphemes, such as clitics, agreement markers, prepositions, and “relational” items carry grammaticalized functions. These grammaticalized functions are valued by the nouns that follow them. The annexed state is, thus, taken here to be a strategy to code the relationship between a grammaticalized function, and its value.

The analysis explains quite well certain distributions of the annexed and free states. For instance, recall from sections “[Common Syntactic Patterns of the States](#)” and “[Syntactic Micro-variations](#)” that post-verbal S and A arguments, as well as direct objects doubled by a pronominal clitic occur in the annexed state. These are explained as follows. Verbs and other predicates in Berber carry grammaticalized functions; a subject function marked by agreement affixes and an object function marked by accusative pronominal clitics. Post-verbal S and A arguments provide the value of this subject function, while clitic-doubled objects provide the value of the object function. As such both occur in the annexed state. This is illustrated in the following examples. In (36a-36b), the nouns in the annexed state *tqɣift* “girl” and *təmyart* “old woman” provide a value for the subject function marked by, respectively, the feminine circumfix *t-t*, and the subject agreement marker *t-*. In (36c), the

noun *jəḥbubən* “figs” provide a value the object function carried by the pronominal clitic *tnt* “them”.

(36) **Taqbaylit (Mettouchi and Frajzyngier 2013: 18)**

- a. d t-a-gugam-t *t-q/ij-t* nni
 COP FEM-FS-mute-FEM FEM.AS-child-FEM DEM
 “That girl is/was mute.”
- b. ad=ay=idd t-ḥku tə-myar-t
 POT=DAT.1PL=PROX SBJ.3SG.F-tell.AOR F-old.person[ANN]-F.SG
 t-i-muɣufa
 F-ABS.PL-tale.PL
 “The old woman would tell us tales.”
- c. ad=ay=tnt=idd t-əfk
 POT=DAT.1PL=OBJ.3PL.F=PROX SBJ.3SG.F-give.AOR
 tə-myar-t ad = tnt n-əčč
 F-old.person[ANN]-F.SG POT=OBJ.3PL.F SBJ.1PL-eat.AOR
 jə-ḥbub-ən=nni
 ANN.PL.M-dried.figs-PL.M=ANAPH
 “Our grandmother would give them to us to eat, those dried figs.”

The analysis can also easily account for the fact that adjectives in Berber never occur in the annexed state, despite agreeing in gender and number with their nominal head. Mettouchi and Frajzyngier (2013) propose that Berber does not have an attributive function to be valued. And even if Berber had such a function, because adjectives in the language always follow the noun they modify they would not be in the right position to value the attributive function and would not occur in the annexed state.

The account runs into problems, however, when it tries to tackle the relation between states and prepositions. The analysis argues that some prepositions carry a function, which seems to depend on the semantic of the preposition. An instrumental preposition, therefore, carries the function instrumental while another preposition marking direction, for instance, is associated with the directional function. Nominal dependents of such prepositions value these functions and are marked by the annexed state. Other prepositions, particularly those which encode negative semantics, such as *siwa* “except” or *bla* “without” do not carry any function to be valued. Their dependents occur in the free state. It is not clear how the dependent of a preposition could provide a value for a function it carries in the same way that it does with an agreement marker or pronominal clitic. There is some identity relation between the content encoded by agreement markers or pronominal clitics and the nouns in subject or object position. There is no such identity relation between prepositions and their dependents. The various functions that Mettouchi and Frajzyngier (2013) propose can easily be all unified under the concept of dependency. This is the stance taken by Arkadiev (2015) who provides a vehement

criticism of the function-value analysis and develops a strong case for case. The next section focusses on one case account in particular suggesting that the annexed state is an instance of marked nominative case.

Marked Nominative

Within non-Berber scholarship, states are systematically analyzed as cases (Sasse 1984; Aikhenvald 1995; Creissels 2009, 2017; König 2008; Arkadiev 2015). According to these views, the state alternation represents a “poor” case system (following Arkadiev’s (2008) terminology); that is a system only contrasting two cases, one marked by the free state, and the other by the annexed state. Within Berber scholarship, however, two/poor case analyses are, overall, much less popular. Some scholars label states as cases in some of their work but do not elaborate on which particular case these can be associated with (Prasse 1986; Kossmann 1997, 2007, 2013, among others). Mettouchi (2008) proposes a distinction between integrative case (annexed state) and absolute case (free state). Inside noun phrases, the integrative case is an intra-phrasal marker. In clauses, the integrative case is a relation-marking case, linking the nominal in the integrative to the overt or covert argument structure of the predication.

König (2008), based on work from Aikhenvald (1995), analyzes the state alternation as a reflection of a marked nominative case system. Marked-nominative systems are cross-linguistically rare. They are found in several language families of East Africa (e.g., Cushitic, Nilotic), and in one family in southwestern USA (König 2008; Dimmendaal 2014; Handschuh 2014; Baker 2015). A marked nominative system is one in which the form of an intransitive subject (S) is overtly coded and functionally marked, while the case form of an object (O) is functionally unmarked (Dixon 1994; König 2008; Handschuh 2014; Baker 2015). In marked nominative, the form of O arguments has a wider range of distribution and is used, among other things, for citation and left-dislocations (König 2008). Tennet (Nilo-Saharan; South Sudan), a marked-nominative language, illustrates the distribution of nominative and accusative forms in marked nominative systems. In (37), both S (*mányúdí-ɪ* “squirrel”), and the transitive subject (A) (*Lowór-ɪ* “Lowor”), are overtly marked for case by the suffix *ɪ*, but O (*Yomá* “Yoma”) is zero-coded.

(37) Tennet (Surmic, Nilo-Saharan; Randal 1998, in König 2008: 139)

- | | | | |
|----|------------------------|------------------|--------------|
| a. | ók | <i>mányúdí-ɪ</i> | mgínaatí. |
| | go.PFV | squirrel-NOM | there |
| | “Squirrel went there.” | | |
| b. | Ákát | <i>Lowór-ɪ</i> | <i>Yomá.</i> |
| | PFV.spear | Lowor-NOM | Yoma.ACC |
| | “Lowor speared Yoma.” | | |

König (2008) demonstrates that the syntactic distribution of the free state and annexed state mirrors the distribution of accusative and nominative forms in unquestionably marked nominative languages.² While nominative forms are restricted to S and A arguments occurring after the predicate and arguments of prepositions, the accusative form, as shown in (38), is also used for citation (a), nominal predicates (b), S & A occurring in initial positions (c), and a range of peripheral arguments (d).

(38) **Tennet (Randal 1998, cited in König 2006: 538–539)**

- a. kavıyak.
news.ACC
“News.”
- b. k-eénj anná deméz-zóh-t.
1-be 1SG.NOM teach-AGEN.ACC-SG
“I am a teacher.”
- c. lókólı cí á-róh Loham.
Lokuli.ACC AM IMPFV-beat Loham.ACC
“It is Lokuli who is beating Loham.”
- d. k-ε-tééd-a ulúg-t ɔɔ.
1-PFV-cut-1SG fish-SG.ACC head.ACC
“I cut the head of the fish.”

The similarities in distribution between the Berber states and the nominative and accusative forms in marked accusative languages is summarized in the following table.

(39) **Similarities between Marked Nominative Systems and the Berber state alternation**

	Marked Nominative		Berber States	
	Accusative	Nominative	Free State	Annexed State
Unmarked (functionally) form	✓	No	✓	No
Citation form	✓	No	✓	No
S/A in post-verbal position	No	✓	No	✓
S/A in pre-verbal position	✓	No	✓	No
Direct object	✓	No	✓	No
Nominal predicate in copula clauses	✓	No	✓	No
Initial S in copula clauses	✓	–	✓	No
Post-predicate S in copula clauses	–	✓	No	✓
Preposition case	✓	✓	✓	✓

²König (2008) reviews in detail seven languages, most of which have several cases in addition to the nominative and accusative (so the nominative and accusative are part of complex paradigms), and finds that all share some common properties.

From a structural, functional, and typological point of view, the marked nominative case approach is the most explanatory, elegant, and straightforward account of the Berber states. Indeed, like case, the annexed state involves a different coding of nominals, which function as dependents of another head: a verb, a preposition (or an element diachronically developed from a preposition), and arguably another noun. From a typological perspective, the case analysis shares with the construct argument the fact that it links the state alternation to a known coding system, found cross-linguistically in several other languages, including many that are genetically related to Berber. The marked nominative is indeed found in a number of Cushitic languages, such as Somali. Furthermore, as clear from the table in (39), the distributions of the two codings in marked nominative systems fit perfectly with those of the annexed and free states in Berber. The annexed form is not more morphologically marked than the free form, since both states are encoded by prefixes in Berber (Ben Si Said 2014, 2020; Belkadi and Ben Si Said 2017). However, functionally, the free state has the same “freer” distribution as the accusative in marked nominative languages and has, in fact, the distribution cross-linguistically fulfilled by forms assigned to S arguments. The annexed state, like marked nominative forms, occurs in restricted syntactic contexts (cf. sections “[Common Syntactic Patterns of the States](#)” and “[Syntactic Micro-variations](#)”). Interestingly, the case analysis is also the one that fits with most of the configurational models developed to account for the alternation, and described in the section “[Configurational Coding](#).” All these accounts rely either on c-command – a structural relation in which case is often argued to be assigned in Minimalism – or on structural deficiency, another way of representing syntactic dependency. The particular word order in which the annexed state is found, which these analyses have also relied on, can, in fact, be easily explained without compromising a case analysis: arguments, unless dislocated for pragmatic purposes, follow their heads in Berber.

There are, however, several important ways in which the states do not behave canonically as cases. These non-canonical behaviors are hard to ignore when one has a profound knowledge of Berber morphosyntax and the details of the states distribution. These explain why the case analysis is not as popular in Berber scholarship. In the next section, some important objections to the case analysis are discussed and evaluated from a typological perspective.

Challenges to Case in Berber and Typology

In the last section, I showed that the cross-linguistic phenomenon with which the Berber nominal state marking shares most common properties is the Marked Nominative case system, as proposed by Aikhenvald (1995) and most recently by König (2006, 2008). (In this section, I will use the terms “annexed state” and “nominative” as well as “free state” and “accusative” interchangeably.) There are three important challenges that the Berber nominal states pose to this Marked Nominative account: (i) post-verbal S and A arguments occur in the free state/accusative with some quantifiers, (ii) clitic-doubled O surface in the annexed/marked nominative, and

(iii) the lack of state/case agreement on attributive adjectives. The aim of this section is to address these. Specifically, it looks at how typological research on case and descriptive work on other languages can inform us, and it evaluates how problematic these challenges actually are.

Post-Verbal S and A with Quantifiers

In all Berber languages that have the alternation, the nominal arguments of a head governing the annexed state can occur in the free state if they are directly preceded by a member of a small subset of quantifiers – mostly those that do not involve the genitive preposition *n* “of” (cf. section “[Common Syntactic Patterns of the States](#)”). Thus, in the following example from Taqbaylit, the post-verbal A argument *aselmad* “teacher” surfaces in the free state rather than the expected annexed state/nominative, because it is preceded by the quantifier *kul* “every”.

- (40) *yə-qədəm kul a-səlmad/ *u-selmad* iman=is i Sa'id
 3SGM- every fs-teacher as-teacher soul=POSS.3SG DAT Sayid
 introduce.PVF
 “Every teacher introduced himself to Sayid.”

The fact that these quantifiers, rather than the verbs, appear to govern the state/case of post-verbal subjects could be a problem for a case analysis and is often used as an argument supporting configurational types of accounts, such as those focusing on c-command relations, or the function-value analysis developed by Mettouchi and Frajzyngier (2013). From a typological perspective, however, these types of data are not as problematic. Indeed, these can simply be explained as instances of a mismatch between the grammatical role of a noun in a clause (or other constituent) and its case. And such mismatches are cross-linguistically extremely frequent, and as will be discussed further in the next subsection may also not be so rare in Berber. In the constructions under discussion here, the mismatch in case is triggered by the quantifiers. Such a phenomenon occurs in several languages, including Finnish and Russian, but also arguably Arabic, Spanish, Romanian, and Southern Italian dialects (Huomo 2018; Ledgeway et al. 2019; Irimia 2020; Corbett pc, Souag pc).

In Finnish, for instance, canonical S and A arguments occur in the nominative form, which is unmarked. This is illustrated in (41).

(41) Finnish (Uralic; Huomo 2018: 425)

Kirja on pöydä-llä.
 book.NOM be.PRS.3SG table-ADE
 “The book is on the table.”

Some S and A arguments, however, occur in the partitive case. Partitive case is driven by a number of factors including the type of verb, the semantic properties of

arguments, and word order. Crucially, here, partitive case assignment is most commonly associated with quantifiers (Huumo 2018). Hence, it is S and A arguments, which are quantified, that occur in the partitive case rather than the nominative case. Thus, in example (42) below, the S argument *auto* “car” is specified by the quantifier *paljon* “a lot of,” and as such ends up being marked by the partitive suffix *-a*.

(42) **Finnish (Uralic; Huumo 2018: 438)**

Paljon	auto-j-a	seiso-o	piha-lla
a.lot.of	car-PL-PAR	stand-PRS.3SG	yard-ADE

“There are a lot of cars standing in the yard.”

In Russian, a similar process occurs. Nominal arguments that are modified by numeral quantifiers surface in the genitive case rather than their regular “structural” cases, such as nominative or accusative (Bošković 2006). This is shown by the data provided in (43). There, the object argument *mašin* “car” occurs in the genitive case, not in the accusative case which is canonically assigned to objects.

(43) **Russian (Slavic; Bošković 2006: 2)**

- | | | | | |
|----|--------|--------|-------|-----------|
| a. | Ivan | kupil | pjat’ | mašin. |
| | Ivan | bought | five | cars(GEN) |
| b. | * Ivan | kupil | pjat’ | mašiny. |
| | Ivan | bought | five | cars(ACC) |

Closer to Berber, in Classical Arabic, the quantifier *kull* “all” obligatorily assigns genitive case to the NP it governs, regardless of the latter’s grammatical function. This is illustrated in (44) below.

(44) **Classical Arabic (Souag, pc)**

kull-u	n-naas-i
all-NOM	DEF-people-GEN

“All the people.”

There are some differences between the cases discussed here and the Berber “quantifier-induced” free state. First, in Finnish and Russian, but also in the other languages mentioned above (i.e., Spanish, Romanian, and Southern Italian dialects), it is a quantifier, or the semantics associated with a quantified noun phrase, that is directly responsible for the change in case assignment. For instance, Ledgeway et al. (2019) discuss some examples of specially coded quantified object arguments in Southern Italian dialects and show that it is the agentivity of such arguments, which are more likely to be quantified, that triggers the special marking. For Berber, it is not clear whether the quantifiers that trigger free state marking or the arguments they modify share any special semantic or lexical property. Second, the non-canonical

case form found on S and A arguments after some quantifiers in Berber is the unmarked one (i.e., the free state), while in other languages, the special coding linked to quantifiers involves marked or peripheral cases, such as the marked accusative, genitive, or the partitive. There might be a range of explanations for the mismatch of case on S and A arguments, modified by certain quantifiers in Berber, involving diachrony and language change, contact and/or semantics. However, the lack of pan-Berber study of quantifiers and their associated state makes it difficult to verify. The bottom line here is that the morphosyntactic mechanism involved in the Finnish and Russian examples are very comparable to that in Berber.

O Arguments in the Annexed/Nominative Form

As discussed in the section “[Syntactic Micro-variations](#),” several varieties of Taqbaylit and Tamazight allow O arguments to surface in the annexed/ nominative form (Chaker 1988; Bendjaballah and Haiden 2005; Achab 2006; Guerssel 2015). Such arguments can, in principle, co-occur with A arguments, which are then marked in the canonical way; that is in the annexed state/nominative. This is illustrated in (45) below, where the O argument *uyifk* “milk” is in the annexed state, whether the A argument, *wəmcic* “cat” (in the annexed form), is realized or not.

(45) Taqbaylit (Belkadi and Ben Si Said 2017)

y-swa=t	u-yifk	nni	(w-əmcic nni)
3SGM-drink.PFV=3SGM.ACC	AS-milk	DEM	(AS-cat DEM)
“He/The cat drank the milk.”			

Annexed state/nominative O arguments appear to be quite rare and have a special relationship with accusative clitic. Thus, an O argument obligatorily surfaces in the annexed state/nominative if doubled or cross-referenced by an accusative pronominal clitic on the verb, or the functional head directly preceding the verb. While an accusative pronominal clitic can occur without a corresponding O, an O in the annexed state/nominative cannot occur without a clitic. These properties are shown in the following examples.

(46) Taqbaylit (Belkadi and Ben Si Said 2017)

- | | | | |
|----|-------------------------|---------------|-----|
| a. | *y-swa=t | a-yifk | nni |
| | 3SGM-drink.PFV=3SGM.ACC | FS.SG-milk | DEM |
| | “He drank the milk.” | | |
| b. | *y-swa=∅ | u-yifk | nni |
| | 3SGM-drink.PFV | AS-milk | DEM |
| | “He drank the milk.” | | |
| c. | y-swa=t | | |
| | 3SGM-drink.PFV=3SGM.ACC | | |
| | “He drank it.” | | |

(48) **Taqbaylit (Author's own corpus)**

- a. lawan n zit, a n-ruḥ a n-lqəd a-zəmur
 time of oil IRR 1PL-go.AOR IRR 1PL-pick.up.AOR FS-olive
 "At the time of oil, we would go to pick-up olives."
- b. a n-ruḥ a n-lqəd a-zəmur nni
 IRR 1PL-go.AOR IRR 1PL-pick.up.AOR FS-olive DEM
 "We will go to pick up these olives."

(49) **Taqbaylit (Author's own corpus)**

- a. *lawan n zit, a n-ruḥ a=t n-lqəd u-zəmur
 time of oil IRR 1PL-go.AOR IRR=3SGM.ACC 1PL-pick.up.AOR AS-olive
 "At the time of oil, we would go to pick-up olives."
- b. a n-ruḥ a=t n-lqəd u-zəmur nni
 IRR 1PL-go.AOR IRR=3SGM.O 1PL-pick.up.AOR AS-olive DEM
 "We will go to pick-up these olives."

Cross-linguistically, topic arguments cannot generally occur in narrow focus (cf. Lambrecht 1994; Nikolaeva 2001). Interestingly, Taqbaylit O arguments in the free state can be in narrow focus, even when they are not definite (e.g., (50)). However, those marked by the annexed state cannot be in narrow focus, even if they are definite and activated (e.g., (51) and (52)).

- (50) Q. dacu i tuyəḍ?
 What did you buy?
 A. uyə-y=d *taqəndurt.*
 buy.PFV-1SG=VEN dress.FS
 "I bought a dress."
- (51) Q. dacu i tččəmt?
 What did you eat?
 A. ?n-čča=*tt* *cərba* i t-ḍəbwə-ḍ
 1SG-eat.PFV=3SGF.ACC soup.AS REL 2SG-cook.PFV-2SG
 "We ate the soup that you cooked."
- (52) Q. anta cərba i tččəmt?
 Which soup did you eat?
 A. ?n-čča=*tt* *cərba* i t-ḍəbwə-ḍ
 1SG-eat.PFV=3SGF.ACC soup.AS REL 2SG-cook.PFV-2SG
 "We ate the soup that you cooked."

Right-dislocated O arguments, on the other hand, have been shown to be re-activated topics, selected topics from several possible ones (Mettouchi 2011; Lafkioui 2000, 2010, 2014), and after-thoughts (Belkadi and Ben Si Said 2017).

Hence, the annexed state/nominative is here used to code the prominence of a topical object. While prominent O arguments are often argued to be specially coded

in order to be distinguished from A arguments (Aissen 2003; Malchukov 2008; Arkadiev 2008), there is evidence from a number of languages that special coding of O arguments may, in fact, occur in order to mark the strong similarities between them and A arguments (Dalrymple and Nikolaeva 2011). In the following example from Tundra Yukaghir, discussed by Matić (2019), some object arguments, pronouns in certain persons and with a discourse prominent function, surface in the nominative case (53) or the marked nominative case (54).

(53) **Tundra Yukaghir (Matić 2019: 343)**

tude nime-Ø öge-tə-l'al-mələ.
 3SG yurt-NOM stand-CAUS-EVID-SO.3
 "He set up his yurt." (MNT_evil.man&fire_011)

(54) **Tundra Yukaghir (Matić 2019: 346, 357)**

- a. (...) *mit-u-l* jaw-n-o: kewrəj-k.
 1PL-0-NOM all-O take.away-IMP
 "(...) take us all with you." (K2005: T6.155)
- b. e, l'e: *tude- l* par-məN.
 INTJ PTL 3SG-NOM cook-SO.1/2SG
 "I've cooked her." (M2001: T1.250)

This instance of differential case marking, whose purpose is to mark the similarity between a prominent object and a prototypical A argument, gives rise to an atypical type of (contextual) case syncretism. Contextual case syncretism is a morphological phenomenon whereby a case category is expressed by another case's morpheme, given a particular context (Baerman et al. 2005; Calabrese 2008). In Hindi, for instance, as illustrated in (55) and (56) below, the ergative case is realized by the exponent of the dative case on certain A arguments, while the accusative is realized as nominative with certain O arguments.

(55) **Hindi (Indo-Aryan; Mohanan 1990: 104)**

Ilaa-ne haar uTaayaa
 Ila-ERG necklace lift-PFV
 "Ila lifted a/the necklace."

(56) **Hindi (Indo-Aryan; De Hoop and Narasimhan 2005: 64)**

- a. Raam=ko ek kitaab mil-ii
 Raam=DAT one book receive-PFV.SG.F
 "Raam received a book."
- b. Ilaa-ne bacce-ko uTaayaa
 Ila-ERG child-ACC lift.PFV
 "Ila lifted the child."

Cross-linguistically, the most common types of syncretism triggered by differential case marking are those in which a marked case, such as ergative or accusative, is realized by the case form assigned to the S argument (prototypically unmarked), such as the absolutive or nominative, or by a peripheral/oblique case, for instance, dative or genitive (Bossong 1991; Baerman et al. 2005). There are, in fact, direct correlations between differential patterns and the degree of prominence of the affected argument (Dixon 1994; Bossong 1991; Aissen 1999; Malchukov 2007; Dalrymple and Nikolaeva 2011). Thus, an accusative O argument generally surfaces in the nominative case if it is less prominent on the hierarchies discussed above. An accusative O argument is realized in the dative or genitive case if it is more prominent on the hierarchies. To my knowledge, prominent O arguments are never marked by (marked) cases given to A arguments, such as, for instance, the ergative. If the annexed state is indeed an exponent of marked nominative, then what we have here is a phenomenon that does not seem to be exhibited by any other language: an instance of a prominent O occurring in the S and marked A case. In Belkadi (2018), I attempted to explain these observed atypical patterns. Assuming that the state alternation is indeed a case alternation following a marked nominative pattern, I proposed that the atypical marking is triggered by the markedness reversal of the Berber system. In a language where accusative is the unmarked form and nominative is the marked form, it is not surprising to find ACC O = NOM O in reverse contexts too.

The second challenge posed by annexed state/nominative O is the systematic co-occurrence of annexed state O with accusative clitics. Across languages, the two most common strategies for marking topical O arguments are case and agreement or cross-reference (Nikolaeva 2001; Bickel and Nichols 2012; Iemmolo 2010; Dalrymple and Nikolaeva 2011; Seržant and Witzlack-Makarevich 2018). Many languages, in fact, mark prominent O by both differential case marking, and cross-reference/agreement. North-Eastern Neo-Aramaic dialects (Coghill 2014), Spanish and Italian (Leonetti 2004; Estigarribia 2013; Iemmolo and Klump 2014), and Belhare, a Tibeto-Burman language from Nepal (Bickel and Nichols 2012) are, for instance, known to mark special O arguments by both agreement and case. This is illustrated in (57), where the object function is marked by both case on the definite noun *pitcha-chi* “children” and agreement on the verb.

(57) **Belhare (Tibeto-Burman; Bickel and Nichols 2012)**

kubaN-chi-Na	pitcha- chi	n-ten-he- chi
monkey-NSG-ERG	child-NSG-ABS	3NSG.A-hit-PST-3NS.P

“The monkeys hit the children.”

Adjectives and Lack of Case Agreement

Another counterargument to the case analysis, often raised in the Berber literature (Bendjaballah and Haiden 2005; Mettouchi and Frajzyngier 2013) is the lack of case agreement on attributive adjectives. Attributive adjectives in Berber, indeed,

obligatorily agree in gender and number with the noun they modify but always surface in the free state. In (58), the masculine singular noun *argaz* “man” is in the annexed state form (*urgaz*) because it is a dependent of the genitive preposition *n* “of”. The adjective *ameqqran* “tall,” which modifies it, is also in the masculine plural form. However, it is in the free state. In (59), the adjective *taməcɬuht* “little” agrees with the feminine singular gender of *təqcict* “girl,” but again, not with its annexed state.

(58) Figuig Berber (Morocco; Kossmann 1997: 242)

n	u-rgaz	<i>a-meqqran</i>
of	AS-man.SG.MS	FS-tall-SG.MS

“Of the tall man.”

(59) Taqbaylit (Author’s own corpus)

yiwət	<i>təqcict</i>	<i>taməcɬuht</i>
one	girl.AS.FM	little.FS.FM

“A little girl.”

This absence of “state” agreement, or concord as it is sometimes referred to, is perhaps the most problematic, from a typological perspective, of the three challenges addressed in this section. Indeed, if case concord alone is frequently optional (e.g., Nez Perce, Norris 2014) and often occurs without gender or number concord (e.g., Georgian, Norris 2017), the expression of gender and/or number concord without case concord is quite rare cross-linguistically (Rießler 2016; Norris 2017; Baerman pc). In a typological review of concord, Norris (2017: 8) finds only very few languages showing this pattern. Furthermore, all these languages have cases that can, in fact, be analyzed as adpositions. Hebrew, for example, is cited in this study as a language which possibly exhibits definiteness, gender, and number concord without case concord. Thus, in (60a) below, the Hebrew counterpart of the adjective “nice” agrees with the definiteness, masculine gender, and plural number of the noun *hajeladim* “boys,” and surfaces as *hanexmadim*. In (60b), it agrees with the feminine gender of the noun *hajeladot* “girls,” and surfaces as *hanexmadot*. In both examples, the adjective does not agree with the accusative case of the noun modified, marked by the form *et*.

(60) Hebrew (Semitic; Norris 2017: 8)

- | | | | | |
|----|-------------------------|-----------|-----------------|----------------------|
| a. | raiti | et | ha-jeladim | <i>ha-nexmad-im.</i> |
| | I.saw | ACC | the-boys(M.PL) | the-nice-M.PL |
| | “I saw the nice boys.” | | | |
| b. | raiti | et | ha-jeladot | <i>ha-nexmad-ot.</i> |
| | I.saw | ACC | the-girls(F.PL) | the-nice-F.PL |
| | “I saw the nice girls.” | | | |

Swedish (Johan Schalin pc) is also a language which can be argued to have number and gender concord on adjectives without case concord. In the following examples, thus, the forms of the adjective *gott* “good” changes depending on the gender and number of the noun it modifies, but never agrees with the genitive case of the noun.

(61) **Swedish (Germanic; Schalin pc)**

- | | | | |
|----|-----------------------|--------------------|------------------|
| a. | Ett | <i>gott</i> | vin |
| | A | good.NEUT.SG | wine.NEUT.SG |
| | “A good wine.” | | |
| b. | ett | <i>gott</i> | vin-s |
| | A | good.NEUT.SG | wine.NEUT.SG-GEN |
| | “A good wine.” | | |
| c. | Flera | <i>goda</i> | viner |
| | several | good.NEUT.PL | wine.NEUT.PL |
| | “Several good wines.” | | |
| d. | Flera | <i>goda</i> | viner-s |
| | Several | good.NEUT.PL | wine.NEUT.PL-GEN |
| | “Several good wines.” | | |

As with the Hebrew accusative marker, the Swedish genitive morpheme *-s* is often argued to be an adposition, rather than truly affixal.

Several other languages which have morphemes that are unarguably case markers have gender and number agreement on adjectives without case agreement, but only in very specific contexts and in particular ways. Rießler (2016) reviews the typological properties of attributive constructions across hundreds of languages from various linguistic families and different regions of the world. His data show that a more common phenomenon is for adjectives to have defective case paradigms. In some languages, like Chechen, Svan (a Kartvelian language spoken in Georgia) or some Romani dialects, adjectives only mark a distinction between nominative and all other cases. Example (62) illustrates the defective paradigm of adjectives in Chechen.

(62) **Chechen (Nakh-Daghestanian; Nichols 1994: 29, cited in Rießler 2016)**

- | | | | | |
|----|--------------------------|--------------------|---------------|----------|
| a. | <i>dika</i> ⁿ | stag ³ | “good person” | [nom:sg] |
| b. | <i>dikaču</i> | stega ⁿ | | [gen:sg] |
| c. | <i>dikaču</i> | stagana | | [dat:sg] |
| d. | <i>dikaču</i> | staga | | [erg:sg] |
| | (...) | | | |

Defective case marking can also function in slightly more complex ways. Thus, in Ngayasan (Samoyedic), adjectives show agreement with the nominative, accusative, and genitive cases of the noun they modify, but not with other case forms. Hence, when the head noun is in another case, the adjective automatically surfaces in the

genitive. In Itelmen, a Chukotko-Kamchatkan language spoken in Eastern Russia, adjectives only agree with one case: the instrumental case.

A possible example of a language that does not show agreement in case at all, despite having agreement in number is Orok, a Tungusic language spoken in Russia. This is illustrated in (63), where the form of the adjective varies depending on the number of its head noun but not its case.

(63) Orok (Tungusic; Petrova 1967 : 55, cited in Rießler 2016)

- | | | |
|----|--|--------------|
| a. | dāi | dalun) |
| | big | store |
| | “big store (i.e., warehouse, storehouse).” | |
| b. | dāi-l | dalul |
| | big-PL | store-PL |
| | “big stores.” | |
| c. | dāi-l | dalul-tai |
| | big-PL | store-PL-LOC |
| | “in big stores.” | |

In light of this typological data on case agreement and attributive adjectives, three possible hypothetical avenues may be followed on the relationship between the Berber states and case. The first hypothesis is that the states are indeed cases but adjectives in Berber have defective paradigms and only inflect for the free state/accusative. The second hypothesis is that adjectives do not agree in case in Berber, and thus surface in the unmarked form, the free state/accusative. The third hypothesis is that Berber states are not on a par with canonical case markers, but are more alike the genitive and accusative adpositions of Swedish, Hebrew, and other languages surveyed by Norris (2017).

Disappointingly, it is difficult to know which of these three hypotheses is correct. Yet, what is interesting is that the cross-linguistic facts reviewed in this section do not actually rule out the case analysis of the Berber state system. For even in languages like Swedish and Hebrew, the ambiguity of these morphemes can be taken to be a sign of them being somewhere on the grammaticalization cline between adpositions and “true” or “fully-fledged” case markers. And while the two categories of adpositions and case affixes differ from a morphological perspective, functionally, the two can be argued to be akin (Blake 2001). The preposition *a* found in varieties of Spanish and several Italo-Romance languages, for instance, is generally considered to encode case, either accusative, or dative (Fàbregas 2013; Ledgeway et al. 2019). These languages are also always included in typological work on differential case marking. Now, the Berber states can definitely not be analyzed as adpositions. As discussed in sections “The Morphosyntax of the Alternation: Diachrony, Synchrony, and Micro-variations,” these affixes co-occur with prepositions and are more generally argued to be grammaticalized from demonstrative articles. Yet, it is plausible that, like adpositions in Hebrew, Swedish, and the other languages surveyed by Norris (2017), the Berber state affixes are

stuck somewhere on the grammaticalization cline between demonstratives and fully-fledged case markers. If so, the lack of case/state concord on adjectives in Berber shows us that a distinction is needed between the functional nature of the states and their morphological status on the one hand, and their degree of grammaticalization, on the other.

Conclusion

Despite sharing similarities with several well-known morphosyntactic phenomena, such as case, construct state or anti-construct state, the Berber state alternation is clearly difficult to tackle. The many ways in which the alternation differs from prototypical elements in these constructions explain the lack of consensus on what the alternation represents, and its analysis as a new, unrecognized, and (so far) unique phenomenon. In this chapter, I argued that the marked nominative case analysis, proposed by Aikhenvald (1995) and developed further by König (2006), is, at this time, the most elegant and straightforward categorization of the alternation. Indeed, the state alternation shows more similarities to marked nominative systems than to any other cross-linguistic morphosyntactic phenomenon, which cannot be accidental or coincidental. It is clear that in diachrony, the state affixes were demonstrative markers which over time grammaticalized and attached onto nominal stems. There is also evidence, thanks to remnants in some languages, that these elements must have occurred across Berber with prepositions, and in noun phrases. In synchrony, it is clear that, where they are still productive, they participate in coding grammatical functions in the clausal domain too – a canonical S/A argument is in the annexed state and a canonical object is in the free state.

Nonetheless, it is undeniable that the Berber case system is non-canonical in several ways. Berber appears to be one of few languages in the world which display number and gender agreement on attributive adjectives but not case agreement. Like all case systems, it involves differential markings of arguments triggered by the context, such as whether a noun is quantified by certain elements or not, its prominence, and category. This chapter has demonstrated that these differential markings are often the reverse of known typological patterns. These reverse patterns can come from the nature of the system itself: (i) it is “poor,” involving an opposition between two forms only; (ii) the S/A case, the annexed state/nominative, is the functionally dependent/marked case, so is used in differential contexts, where other languages use O or peripheral cases instead.

The rationale behind this chapter was to look at what typological and cross-linguistic research on different systems of nominal coding could tell us about the nature and function of the Berber state alternation. As I reach the conclusion, there are many questions and issues that I have no choice but to leave unanswered and unaddressed. As descriptive, variationist, dialectological, and diachronic research on Berber languages develops, I hope that we can uncover data and find explanations for these remaining puzzles.

References

- Achab K (2003) Alternation of state in Berber. In: Lecarme J (ed) *Research in Afroasiatic grammar II. Current issues in linguistic theory*. John Benjamins Publishing, Amsterdam, pp 1–20
- Achab K (2006) Internal structure of verb meaning: a study of verbs of (change of) state in Tamazight (Berber). PhD dissertation. University of Ottawa. UMI, Ann Arbor
- Aikhenvald A (1995) Split ergativity in Berber languages. *St Petersburg J Afr Stud* 4:39–68
- Aissen J (1999) Markedness and subject choice in optimality theory. In: *Natural Language & Linguistic Theory* 17, pp 673–711
- Aissen J (2003) Differential object marking: iconicity vs. economy. *Nat Lang Linguist Theory* 21: 435–483
- Applegate JR (1970) The Berber languages. In: Ferguson CA (ed) *Linguistics in south west Asia and North Africa*. Mouton, Berlin/Boston, pp 586–662
- Arkadiev PM (2008) Poor (two-term) case systems: limits of neutralization. In: Malchukov AJ, Spencer A (eds) *The Oxford handbook of case*. Oxford University Press, Oxford
- Arkadiev P (2015) The Berber “state” distinction: dependent marking after all? A commentary on Mettouchi & Frajzyngier (2013). *Linguist Typol* 19(1):87–110
- Baerman M, Brown D, Corbett GG (2005) *The syntax-morphology interface: a study of syncretism*. Cambridge University Press, Cambridge, UK
- Baker M (2015) *Case: its principles and its parameters*. Cambridge studies in linguistics. Cambridge University Press, Cambridge
- Basset A (1932) Note sur l'état d'annexion en berbère. *Bulletin de la Société de Linguistique de Paris* 33:173–174
- Basset A (1952) *La langue berbère*. Oxford University Press, London
- Basset A, Picard A (1948) *Éléments de grammaire berbère (Kabylie-Irdjen)*. Éditions “La Typo-Litho” et J. Carbonel réunie, Alger
- Belkadi A (2018) Syncretism in a marked nominative language: case and DOM in Taqbaylit. Poster presented at the Annual Meeting of the Linguistic Association of Great Britain at the University of Sheffield (UK). Sept 11th–14th 2018
- Belkadi A, Ben Si Said S (2017) Marked nominative in Berber? Views from the phonology-syntax-information structure Interface. Paper presented at the 45th North Atlantic conference on Afro-Asiatic linguistics. University of Leiden (Leiden). June 9th–11th 2017
- Ben Si Said S (2014) *De la nature de la variation diatopique en kabyle: étude de la formation des singulier et pluriel nominaux*. Unpublished PhD dissertation, Université de Nice
- Ben Si Said S (2020) The initial vowel of nouns and the construct state in Kabyle (Berber). *Can J Linguist/Revue Canadienne de Linguistique* 65:155–180
- Bendjaballah S, Haiden M (2005) The grammar of prepositions in Berber. Ms. <https://ling.auf.net/lingbuzz/000261/>. Last accessed 20 Apr 2023
- Benmamoun E (2000) *The feature structure of functional categories: a comparative study of Arabic dialects*. Oxford University Press, New York/Oxford
- Bickel B, Nichols J (2012) Inflectional morphology. In: Shopen T (ed) *Language typology and syntactic description*. Cambridge University Press, Cambridge, pp 169–240
- Blake BJ (2001) *Case*. Cambridge textbooks in linguistics, 2nd edn. Cambridge University Press, Cambridge
- Bošković Ž (2006) Case and agreement with genitive of quantification in Russian. In: Boeckx C (ed) *Agreement systems*. John Benjamins Publishing, Amsterdam, pp 99–120. [Linguistik Aktuell/Linguistics Today 92]
- Bossong G (1991) Differential object marking in romance and beyond. In: Wanner D, Kibbee DA (eds) *New analyses in romance linguistics. Selected papers from the linguistic symposium on romance languages XVIII, Urbana-Champaign, April 7–9, 1988*. John Benjamins Publishing, Amsterdam, pp 143–170
- Brugnatelli V (1987) Deux notes sur l'état d'annexion en berbère. In: Jungraithmayr-W H, Müller W (eds) *Proceedings of the 4th int. Hamito-Semitic congress*. John Benjamins Publishing, Amsterdam/Philadelphia, pp 349–359

- Calabrese A (2008) On absolute and contextual syncretism: remarks on the structure of case paradigms and on how to derive them. In: Bachrach A, Nevins A (eds) *Inflectional identity*. Oxford University Press, Oxford, pp 156–205
- Chaker S (1983) *Un parler berbère d'Algérie (Kabylie): Syntaxe*. Publications de l'Université de Provence, Aix-En-Provence
- Chaker S (1988) Annexion (état d'). In: Chaker S (ed) *Encyclopédie berbère*. Peeters Publishers, Leuven
- Chaker S (1998) Fonctions Syntaxiques. In: Chaker S (ed) *Encyclopédie berbère*. Peeters Publishers, Leuven
- Coghill E (2014) Differential object marking in Neo-Aramaic. *Linguistics* 52(2):335–364
- Creissels D (2009) Construct forms of nouns in African languages. In: Austin PK, Bond O, Charette M, Nathan D, Sells P (eds) *Proceedings of conference on language documentation & linguistic theory 2*. SOAS, London, November 13–14, 2009, pp 73–82
- Creissels D (2017) Construct forms of nouns in typological perspective. Paper presented at the 50th annual meeting of the Societas Linguistica Europea. Zurich University (Zurich). 10–13 September 2017
- Dalrymple M, Nikolaeva I (2011) *Objects and information structure*. Cambridge studies in linguistics. Cambridge University Press, Cambridge
- De Hoop H, Narasimhan B (2005) Differential case-marking in Hindi. In: Amberber M, de Hoop H (eds) *Competition and variation in natural languages: the case for case*. Elsevier, Oxford, pp 321–346
- Dimmendaal G (2014) Marked nominative systems in Eastern Sudanic and their historical origin. *Afrikanistik Online* 11(3):1–18. <https://www.afrikanistik-aegyptologie-online.de/archiv/2014/3859>. Last accessed 20 Apr 2023
- Dixon RMW (1994) *Ergativity*. Cambridge University Press, Cambridge
- El Hankari A (2014) The construct state in Tarifit Berber. *Lingua* 148:28–52
- El Moujahid H (1993) *Syntaxe du groupe nominal en berbère Tachelhiyt (Parler d'Igherm, Souss, Maroc)*. Thèse de Doctorat d'état ès-Lettres. Rabat: FLSH, Université Mohammed V
- El Moujahid El H (1997) *Grammaire générative du berbère: Morphologie et Syntaxe du Nom en Tashelhit*. Série Thèses et mémoires n. 38. Université Mohamed V
- Engelhardt M (2000) The projection of argument-taking nominals. *Nat Lang Linguist Theory* 18: 41–88
- Ennaji M (2001) The construct state in Berber. In: *Studies in the linguistic sciences*, vol 31, no 2. University of Illinois Publication, Urbana-Champaign, pp 55–72. <https://www.ideals.illinois.edu/items/9591#:~:text=http%3A/hdl.handle.net/2142/9551>. Last accessed 20 Apr 2023
- Estigarribia B (2013) Rioplatense Spanish clitic doubling and “tripling” in lexical-functional grammar. In: Howe C, Blackwell SA, Lubbers Quesada M (eds) *Selected proceedings of the 2011 Hispanic linguistics symposium*. Cascadia Proceedings Project, Somerville, pp 297–309
- Fàbregas A (2013) Differential object marking in Spanish: state of the art. *Borealis Int J Hispanic Linguist* 2(2):1–80
- Fassi-Fehri A (1999) Arabic modifying adjectives and DP structures. *Studia Linguistica* 53: 105–154
- Galand L (1964) L'énoncé verbal en berbère. *Étude de fonctions*. Cahiers Ferdinand de Saussure 21: 33–53
- Galand L (1969) Types d'expansion nominale en berbère. Cahiers Ferdinand de Saussure 24: 83–100
- Guerssel M (1992) On the case system of Berber. *Can J Linguist/Revue Canadienne De Linguistique* 37(2):175–196
- Guerssel M (1995) Berber clitic doubling and syntactic extraction. *Can J Linguist/Revue Canadienne De Linguistique* 24(1):111–133
- Handschuh C (2014) A typology of marked-S languages. *Studies in diversity linguistics* 1. Language Science Press, Berlin
- Heath J (2006) *A grammar of Tamashek (Tuareg of Mali)*. Mouton, Berlin/New York

- Huumo T (2018) The partitive A: on uses of the Finnish partitive subject in transitive clauses. In: Seržant IA, Witzlack-Makarevich A (eds) *Diachrony of differential argument marking. Studies in diversity linguistics 19*. Language Science Press, Berlin, pp 423–453
- Iemmolo G (2010) Topicality and differential object marking: evidence from romance and beyond. *Stud Lang* 34(2):239–272
- Iemmolo G, Klumpp G (2014) Introduction. In: Iemmolo G, Klumpp G (eds) *Special issue in differential object marking: theoretical and empirical issues. Linguistics* 52(2):271–279
- Irimia MA (2020) Variation in differential object marking: on some differences between Spanish and Romanian. *Open Linguist* 6:424–462
- K'halainou M (2006) Tajerrunt n Tmazitt (Tumzabt). *Grammaire Berbère. Le Mozabite. Ghardaia*
- König C (2006) Marked nominative in Africa. *Stud Lang* 30(4):655–732
- König C (2008) *Case in Africa*. Oxford University Press, Oxford
- Kossmann MG (1997) *Grammaire du parler berbère de Figuig (Maroc oriental)*. Ussun amazy 10. Peeters, Leuven
- Kossmann MG (2000) *Esquisse Grammaticale du Rifain Oriental*. Ussun amazy 16. Peeters, Leuven
- Kossmann M (2002) Person-number inflection in Berber verbs. In: *Sprawozdania z posiedzen komisji naukowych (Polska Akademia Nauk, oddzial w Krakowie)* 44(2):59–62
- Kossmann MG (2007) Berber morphology. In: Kaye AS (ed) *Morphology of Asia and Africa*. Eisenbrauns, Winona Lake, pp 135–145
- Kossmann MG (2013) *A grammatical sketch of Ghadames Berber (Libya)*. Berber studies 40. Rüdiger Köppe Verlag, Köln
- Lafkioui M (2000) *Syntaxe intégrée de l'énoncé non-verbal rifain: l'énoncé auxiliaire de prédication spécifique*. *Comptes rendus du G.L.E.C.S.* XXXIII:165–187
- Lafkioui M (2010) La topicalisation en berbère: formes et structures. In: Stroomer HM, Kossmann MG, Ibrizimow D, Vossen R (eds) *Etudes berberes V Essais sur des variations dialectales et autres articles*. Berber Studies, vol 28. Rüdiger Köppe Verlag, Köln, pp 121–132
- Lafkioui M (2014) Topicalization in Berber: a typological perspective. *STUF Lang Typol Univ* 67(1):97–112
- Lahrouchi M (2011) Têtes et gabarits. *Etudes et Documents Berbères* 29/30:255–274
- Lahrouchi M (2013) Templates, markers and syntactic structure in Tashlhiyt Berber. *Lingua* 133: 53–72
- Lambrecht K (1994) *Information structure and sentence form: topic, focus, and the mental representations of discourse referents*. Cambridge studies in linguistics. Cambridge University Press, Cambridge
- Lanfry J (1971–1972) Deux notes sur le berbère de Ghadamès. *Comptes rendus du Groupe Linguistique d'Études chamito-sémitiques G.L.E.C.S.* XVI:175–184
- Laoust E (1912) *Etude sur le dialecte du Chenoua comparé avec des beni-menacer et des beni-salah*. In: *Bulletin de correspondance africaine*, 50. Publications de la Faculté des Lettres d'Alger. Ernest Leroux, Paris
- Ledgeway A, Schifano N, Silvestri G (2019) Differential Object Marking and the properties of D in the dialects of the extreme south of Italy. *Glossa* 4(1):1–25. <https://doi.org/10.5334/gjgl.569>
- Leonetti M (2004) Specificity and differential object marking in Spanish. *Catalan J Linguist* 3: 75–114
- Mahootian S (1997) *Persian*. Routledge, London
- Malchukov A (2007) Animacy and asymmetries in differential case marking. *Lingua* 118:203–221
- Malchukov AL (2008) Animacy and asymmetries in differential case marking. *Lingua* 118(2): 203–221
- Matić D (2019) A case for syntactic case: the accusative in tundra Yukaghir. *Morphology* 29: 337–383
- Mettouchi A (2008) Case-marking, syntactic domains and information structure in Kabyle (Berber). In: Frajzyngier Z, Shay E (eds) *Interaction of morphology and syntax: case studies in Afroasiatic*. John Benjamins Publishing, Amsterdam

- Mettouchi A (2011) The grammaticalization of directional clitics in Berber. Paper presented at the 44th meeting of the Societas Linguistica Europaea (SLE) workshop “come and go off the grammaticalization path”, Sep 2011, Logrono, Spain
- Mettouchi A (2014) Foundations for a typology of the annexed/absolute state systems in Berber. *Stuf* 67(1): 47–61. Mouton De Gruyter
- Mettouchi A, Frajzyngier Z (2013) A previously unrecognized typological category: the state distinction in Kabyle (Berber). *Linguist Typol* 17:1–30
- Mohanan T (1990) Arguments in Hindi. Doctoral dissertation, Stanford University, Stanford, California
- Mourigh K (2015) A grammar of Ghomara Berber. Doctoral dissertation, Rijksuniversiteit te Leiden
- Nichols J (1994) Chechen. In: Riexs Smeets (ed) *North East Caucasian Languages Part 2*, Delmar/Caravan Books, New York, pp 3–76
- Nikolaeva I (2001) Secondary topic as a relation in information structure. *Linguistics* 39(1):1–49
- Norris M (2014) A theory of nominal Concord. Unpublished PhD dissertation. UC Santa Cruz
- Norris M (2017) Description and analyses of nominal concord. *Language and linguistic compass*. <https://ling.auf.net/lingbuzz/003739>. Accessed 10 Feb 2020
- Ouhalla J (1993) Subject-extraction, negation and the antiagreement effect. *Nat Lang Linguist Theory* 11:477–518
- Ouhalla J (1996) The construct state in Berber. In: Lecarme J, Lowenstamm J, Shlonsky U (eds) *Studies in Afroasiatic*. John Benjamins Publishing, Amsterdam, pp 278–301
- Penchoen TG (1973) *Tamazight of the Ayt Ndhir*. Afroasiatic dialects, 1. Undena Publishing, Los Angeles
- Petrova TI (1967) *Jazyk orokov (ul'ta)*. Nauka, Leningrad
- Prasse K-G (1974) *Manuel de grammaire touarègue (tahaggart)*, 4–5: nom. Akademisk Forlag, Copenhagen
- Prasse K-G (1986) The values of the tenses in Tuareg (Berber). *Orientalia Suecana*, vol XXXIII–XXXV (1984–1986). Almqvist & Wiksell International, Stockholm
- Randal S (1998) A grammatical sketch of Tenneset. In: Dimmendaal GJ, Last M (eds) *Surmic languages and cultures*, Rüdiger Köppe, Köln, pp 219–272
- Rießler M (2016) *Adjective attribution*. Studies in diversity linguistics 2. Language Science Press, Berlin
- Ritter E (1991) Two functional categories in noun phrases: evidence from modern Hebrew. In: *Syntax and semantics: perspectives on phrase structure: heads and licensing*, vol 25. Academic, London, pp 37–62
- Sasse H-J (1984) Case in Cushitic, Semitic and Berber. In: Bynon J (ed) *Current progress in afroasiatic*. Papers of the third international Hamito-Semitic congress, London 1978. John Benjamins, Amsterdam, pp 111–126
- Seržant IA, Witzlack-Makarevich A (eds) (2018) *Diachrony of differential argument marking*. Studies in diversity linguistics 19. Language Science Press, Berlin
- Souag ML (2010) *Grammatical contact in the Sahara: Arabic, Berber, and Songhay in Tabelbala and Siwa*. Doctoral dissertation, SOAS University of London
- Stumme H (1899) *Handbuch des Schilchischen von Tazerwalt*. J. C. Hinrichs'sche Buchhandlung, Leipzig
- Taifi M (2012) De quelques propositions non verbales en berbère: constructions et effets de sens. In: Ibrizimow D, Vossen R, Stroemer H (eds) *Etudes berbères VI Essais sur la syntaxe et autres articles*. Berber Studies, vol 28. Rüdiger Köppe Verlag, Köln, pp 257–268
- Vycichl W (1957) L'article défini du berbère. In: *Mémorial André Basset (1895–1956)*. Maisonneuve, Paris, pp 139–146

Tashlhiyt Berber DP Assembled

23

Hassan Makhad

Contents

Introduction	494
Nominal Structure in TsB	495
The Internal Structure	495
External Structure	499
The Proposal	501
Initial Vowel Status	501
The DP Organization in TsB	504
Conclusion	513
References	513

Abstract

The DP hypothesis (cf. Abney. The English noun phrase in its sentential aspects. Ph. D. dissertation, MIT, 1987), whereby noun phrases (NPs) are claimed to actually be determiner phrases (DPs), has generated much disagreement in literature. Some researchers took the hypothesis for granted: Longobardi (*Linguist Inq* 25(4): 609–665, 1994), Cinque (On the evidence for partial N-movement in the Romance DP. In: Cinque G, Koster J, Pollock J-Y, Rizzi L, Zanuttini R (eds) *Paths towards universal grammar: study in honor of Richard S. Kayne*. Georgetown University Press, Washington, pp 85–110, 1994), and Progovac (*J Linguist* 34:165–179, 1998). Others have stood against it: Lyons (*Definiteness*. Cambridge University Press,

I would like to thank the editors of the present volume on *Berber Languages and Linguistics*, Professor Karim Bensoukas and Professor Alireza Korangy, for their time, and coordinating efforts to make the project successful. I am extremely grateful to two anonymous reviewers for their valuable comments and thoughtful suggestions concerning the content of this chapter. Their constructive remarks have definitely improved both the quality and readability of my document. Needless to say: all errors are mine.

H. Makhad (✉)

The Polydisciplinary Faculty of Safi, Cadi Ayyad University, Safi, Morocco

e-mail: h.makhad@uca.ma

Cambridge, 1999), Bošković (Left branch extraction, structure of NP, and scrambling. In: Sabel J, Saito M (eds) *The free word order phenomenon: its syntactic sources and diversity*. Mouton, Berlin, pp 13–73, 2005; What will you have, DP or NP? In: E. Elfner, M. Walkow (eds) *Proceedings of the Northeastern Linguistics Society (NELS) 37*. GSLA Publications, Amherst, pp 101–114, 2008). The present work provides an in-depth reconsideration of the nominal construction in Tashlhiyt Berber. I argue that DP projects in the syntax of the language. I present ample empirical arguments that support N merger with little *n*, and with D^0 . The proposed analysis proves to be useful in understanding most of the concealed aspects of the noun and its interaction with pre- and post-nominal modifiers.

Keywords

Tashlhiyt Berber · DP · *nP* · NP · Noun modifiers

Introduction

This chapter is an investigation of the noun form in Tashlhiyt Berber (TsB). It is built on the contemporary and ongoing research on nominal structures in Amazigh/Berber languages. However, this work departs from existing research proposals in a number of ways. With regards to DP in Amazigh studies, there are three trends. The first camp claims that some Amazigh dialects demonstrate the existence of an autonomous D category, as in Ouhalla (1988) and Guerssel (1992). They consider the initial vowel of a noun to be a D-element. For example, they analyze nouns like *argaz* “man” to consist of two morphemes, wherein the initial vowel is a determiner. The second group declines such suggestion. For them the first vowel alternations are manifestations of morphological Case. Hence (in)definiteness resides in discourse outside the domain of syntax (cf. Ouhalla 1996; Achab 2003). The third party considers the first vowel to be a nominalizer unit in the noun category (El Moujahid 1997; Lahrouchi 2013). Thus, *argaz* remains a single unit which may mean “a man” or “the man.”

The objective of this chapter is to further explore the noun category in TsB by considering, both its internal and external properties. This endeavor sheds new light on some overlooked aspects of the noun, together with its interaction with other orbiting elements in its entourage. I argue that the initial vowel forms an integral part of the noun in TsB. It is endowed with a specified gender feature. I assert that phi-features are matched/valued within little *nP* projection. Likewise, I present adequate evidence for the inclusion of periphrastic genitive construction in the system. Furthermore, I present a new account supporting the projection of a DP. I demonstrate that even if TsB does contain overt autonomous articles, it has a minority of D-categories (By D-categories here, I simply mean the elements that may occupy the D head of DP. For example in English, D^0 may host articles, demonstratives or the genitive “’s.”). The reason is that D comprises uninterpretable features and additional information that trigger merger of the head noun, internally. The head of DP is, therefore, the locus of (in)definiteness marks. Definite D^0 forces N-raising, while indefinite one does not do so. The analysis is framed in the Minimalist Program (MP), as in Chomsky (1995, 2001), and subsequent works.

The chapter is organized in two sections. The first explores the morpho-syntactic structure of nominal forms in TsB. It examines the internal and external makeup of the noun category. Section two contains the proposal of the chapter. It initially exposes the nature of the initial nominal vowel and proposes that it is an integral part of the noun which is responsible for masculine gender. After that, it demonstrates the projection of a DP form despite the absence of independent articles in the system. This claim is proven by the syntactic behavior of the noun and its interaction with adjacent modifiers.

Nominal Structure in TsB

Examination of nominal constructions in TsB shows that they delineate internal and external properties. The inner makeup relates to the morphological/morpho-syntactic composition. In contrast, the outer frame represents the syntactic attributes of the declension. The distinction is of particular importance so as to understand the nature of the noun category.

The Internal Structure

This research begins by showing some explicit clarifications about the inner structure of nouns in TsB. Morphologically, TsB nominals differ with regards to number, gender, and state alternations (cf. Boukhris et al. 2008; Lahrouchi 2013; van Putten 2016).

Number

With regard to number, TsB nouns have singular and plural forms. This distinction is, however, not a simple one. Indeed, so many scholars have noticed that Berber number derivation involves various and complex morphological operations (cf. Mourigh 2015). The most dominant one is initial vowel¹ alternation where

¹I would like to specify here that what is referred to as initial vowel may not truly be a mere vowel. Intuitively, it seems that the vowel is preceded by a glottal stop (?-). Thus, the form *argaz* “man” may actually be *?argaz*. This impression appears to be legitimate as it is instantiated in forms like:

- i) qq ?argaz ur-ay i-skar yik
 indeed a-man not-modal he-doing like-this
 “A gentleman would not behave like that.”

If this is true, there is essentially some overlooked equivalence between masculine and feminine in TsB. Gender may be a separate morpheme, as in:

- ii) a. ?a-muɣf b. ta-muɣf
 “cat (masculine)” “cat (feminine)”

Due to space limitations, I leave this issue to further exploration. Hence, I will just follow the trend of beginning masculine nouns with a vowel.

a becomes *i*, coupled with *-n* suffixation, as in *adrar - idrarn* “mountain.” When the form ends in *u* or *i*, the latter change into *a*, as in *aduku - idukan* “slipper; shoe,” and *itri - itran* “star.” If the singular ends in *a*, the plural is suffixed with *-wn*, as in *amksa - imksawn* “shepherd.”

Despite the rich and complex nature of number derivation in Berber, Idrissi (2000) managed to differentiate three types of plurals. The first is internal (aka broken) plurals. They are formed by vowel alternations, as in *amzzuγ - imzzaγ* “ear.” The second type are external plurals. They involve initial vowel change besides an “*-n*” suffixation, as in *afulus - ifullusn* “rooster.” The third class is mixed plurals. They require both vowel alternations and “*-n*” suffixation, as in *aḌaḌ - iḌuḌan* “finger.”²

Moreover, Nejme et al. (2013) as well as Taghbalout et al. (2015) make a four-type model distinction. In addition to the above tripartite division, they ascertain plurals formed by *id-* prefixation, as in *xali - idxali* “my uncles” (cf. Bensoukass 2018). Likewise, Alderete, et al. (2015) distinguish five patterns of plurals. Besides the abovementioned groups, they single out what they call “special plurals”. These include Arabic loan words (*lbħr - lbħur* “sea”) and TɔB words that form plurals with distinct forms, as in *tiTT - alln* “eyes.”

On the basis of these nominal distinctions (Despite these attempts, nominal number derivations sometimes remain unpredictable. This unpredictability results from the complex morphology of the language. The complication of the system is most likely due to historical changes as well as contact with other foreign languages.), number is recognized as a nominal feature. This recognition is essential as number contributes to interpretation, both in syntax and semantics. Both forms are listed in the lexicon with their specified idiosyncratic features.

Gender

Gender is a recognized nominal distinction in TɔB. The language discerns masculine “*awtm*” from feminine “*tawmt*,” morphologically. Generally, masculine

²It is worth pointing out that despite these classifications, there are exceptions, especially when it comes to the alternations of the initial vowel. The idea is that some vowels resist change. Consider this table:

i)	Singular	Plural	Glossary
	anZiḌ	anZiḌ-n	Squirrels
	agar	agar-n	crickets
	aga	agiwn	Buckets for drawing water from a well
	afa	afat-n	Top of a mountain/hill
ii)	imirj	imarijn	Edge of fields in mountain farming
	iyirdm	iyardmiwn	Scorpion
	imillw	imalliwn	Collared Doves
iii)	udm	udmawn	Faces
	ul	ula(wn)	Hearts

nouns begin with a vowel (a-, i-, u-). (However, there are few nouns that begin with consonants, such as *laẓ* “hunger,” *baqzzwa* “bee bird,” *ssṬart* “roof- barrier,” *ḌḌmẓ* “animal trap.”) as in *aggurdj* “flea,” *imi* “door, mouth,” *uffjn* “jackal.” The initial vowel thus indicates masculinity in nouns. (The absence of these initial vowels in some varieties, such as that of Ait Sghrushen (cf. Kossman 2014) may be a mere phonetic deficiency. It seems to me that the initial vowel is present in the base form and its template representation.) The feminine is expressed by the manifestation of the “*t*” morpheme. The latter takes three forms. Normally, it is realized with a circumflex, as in *t-afrux-t* “girl.” Sometimes, it is only a prefix, as in *t-igmmi* “house.” Other times, it is a suffix, as in *lxnf-t* “bag” (cf. Nejme et al. 2013).

Given these facts, it is legitimate to claim that the feminine form is derived from the masculine by *t* affixation (I am aware that sometimes the masculine and the feminine are realized as different words. The feminine of *izimr* “ram” is *tahrujt* “iwi.” Similarly, the feminine of *?ačlluf* “bull” is *tafunast* “cow.” Equally, *ankkur* “buck” has its feminine as *taɣatt* “doe.”). In this sense, the masculine is the default form and the feminine is the marked form, as it incorporates the extra affix *t*. This claim can be supported by the specific fact that gender opposition in TsB sometimes involves size distinction. For instance, the masculine may be used to denote enormity, as in *igmmi* (It is worth noting that the word *tigmmi* “house” is originally feminine, as indicated by the prefix *t*. The removal of this feminine morpheme awards the item *igmmi* an exceeding abnormal size in meaning. This practice is recursive in TsB. For example, *tamsrijt* “room” can be rendered *amsrij* “a large room,” *taddagt* - *addag* “tree,” *timkilt* - *imkil* “bowl,” . . .etc.) “a large house.” At the same time, the feminine can convey smallness and miniaturization of an object (cf. Kossman 2009). For example, *afus* “hand” can receive a diminutive reading by *t* affixation, as in *tafust* “small hand.” Thus, immensity is associated with masculinity, and diminutiveness with femininity.

Moreover, the feminine form may be used to express flattering compliments. For example, *alln* “eyes” has a diminutive form *tiwallin* which is understood as “small and beautiful eyes.” The same is true for the opposition *imi* “a mouth” and *timit* (In some dialects, it is realized as: *timimmit* “a small and beautiful mouth.”) “a small and beautiful mouth.”

Gender distinctions, accordingly, play a potential significant role in characterizing nominal forms. Gender properties are, therefore, relevant for managing syntactic derivations. This relevance is central for deriving adequate structures.

State Alternations

In addition, TsB nouns display state alternations (Actually, this phenomenon is particular to Berber/Amazigh languages. Its appellation may, thus, be inaccurate. The reason is that its observable facts are challenging to adequately understand.). Actually, studies in Amazigh linguistics distinguish two states: Free State (FS) and Construct State (CS) (cf. Guerssel 1983; Ennaji 2001; El Moujahid 1997; Achab 2003; Lahrouchi 2013). The CS is a morpho-syntactic variant form of the

FS. Therefore, the context of the emergence of one negates the presence of the other. The FS *afrux* “boy” is realized as *w/ufrix* in CS. In feminine forms, the initial vowel disappears completely, as in the FS *tafruxt* “girl,” which becomes *tfruxt* in the CS.

Basically, there are four restricted contexts where CS nominals appear. Consider (1):

- 1) a. i-ʃʃ-a urgaz (*urgaz* is actually realized as *wrgaz*.) /*argaz
ayrum
He-eat-Perf. the-man.CS/ the-man.FS bread.FS
“The man ate bread.”
- b. t-dd-a dar wanu / * anu
she-go-perf. near the-well.CS / the-well FS
“She went to the well.”
- c. i-sɣ-a akal n tmzgid / *timzgid
he-buy-perf. the-land of the-mosque CS / the-mosque FS
“He bought the land of/from the mosque.”
- d. j-ʊʃk-a-d urgaz d ufrux /*afrux
he-come-perf.-SP the-man CS and the-boy CS/ the-boy FS
“The man and the boy came.”

On the basis of (1), a TsB noun can be in CS when it is a post verbal subject (1a), object of a preposition (1b), subsequent noun in a genitive construction (1c), or second nominal in a coordination formation (1d). FS forms appear elsewhere.

As a matter of fact, these alternations are markedly confusing in the literature on Berber languages. They are confounded with Case, as in Guerssel (1992), Ouhalla (1996), Achab (2003), Arkadiev (2015), and Felice (2020). The two occurrences are dissimilar though. Actually, TsB makes it obvious that this confusion is improper. One argument that rejects the case analysis of state alternations comes from forms such as:

- 2) a. i-mmudd-a urgaz
he-travel-perf. the-man.CS
“The man travelled.”
- b. argaz i-mmudd-a
the-man.FS he-travel-perf.
“The man travelled.”

The constructions in (2) show that state alternations do not match morphological case realizations in TsB (Bendjaballah and Haiden (2005) have reached the same conclusion in their study of Chemini Berber, a dialect of Taqbaylit. They have established that state alternations and Case are two distinct things.). In (2a), the CS is certainly not linked with nominative case. The reason is that the same subject fails

to retain its case in SVO order, as (2b) illustrates. It is not possible that subjects in VSO order get nominative case, while in SVO they lose that case³ property.

Another inadequacy of associating case with state alternations comes from the inconceivable facts in (1b–1c) where the nouns following the prepositions are interpreted as marked nominative. Objects of prepositions normally bear objective case rather than nominative case. Therefore, either Case is not uniform in TsB, or correlating case with state alternation is wrong. The latter, I believe, is more factual. Therefore, the alternation cannot be accounted for if associated with case.

It accordingly seems that Case and States are two distinct and inherent attributes of TsB. Obviously, Case is not realized morphologically in the language, yet state alternations are a commonplace. Nevertheless, they remain a mysterious facet in TsB, as they do not have an available direct explanation.

As has been illustrated, number, gender, and state alternations are relevant for syntactic derivations. They contribute to the functioning of merger operations and determine relationships among syntactic components so as to ensure convergent derivations. Hence, they are syntactically important, as they provide useful information that is relevant for understanding how nominal forms are construed in TsB.

External Structure

A hallmark of nominal forms in TsB is their ability to co-occur with certain discretionary elements that function, either as pre-modifiers or post-modifiers. These include adjectives, demonstratives (*Dem*), clitics, quantifiers, and numerals. The noun heads the nominal projection and regularly collocates with these other items to form larger nominal units.

As regards word order details, the noun consistently appears before adjectives, as in:

³Note that in languages, such as Literary Arabic, where case is morphological, the case properties are unaffected by the VSO and SVO orders, or any other orders. The subject remains nominative in all orders in (i).

- 1) a. ?akala r-razul-u ?al-yaða:ʔ-a
Ate-he the-man-Nom. The-lunch-Acc.
“The man ate the lunch.”
- b. ?ar-razul-u ?akala ?al-yaða:ʔ-a
the-man-Nom. ate-he The-lunch-Acc.
- c. ?al-yaða:ʔ-a ?akala r-razul-u
The-lunch-Acc ate-he the-man-Nom.
- d. ?al-yaða:ʔ-a ?ar-razul-u ?akala
The-lunch-Acc the-man-Nom. ate-he

- 3) tiggmi imqqurn
house big
“The / a big house”

Moreover, nouns constantly precede demonstratives and clitics, as in:

- 4) a. argaz-ad b. argaz-n-ns
man this man-of-hers
“This man” “Her husband”

However, TsB nouns permanently follow numerals and quantifiers, as in:

- 5) a. jan urgaz b. kra urgaz
one man some man
“A man” “Some man”

Furthermore, it is possible that the two types of modifiers co-occur, simultaneously, in a nominal structure, as in:

- 6) a. jan urgaz igzzuln b. argaz-ad igzzuln c. kullu irgazn-ad gzzulnin
one man short.Sg. man-this short.Sg. all men-these short.PL
“One short man” “This short man” “All these short men”

The form in (6a) shows that a numeral and an adjective may occur together. In (6b), a demonstrative is allowed to appear in tandem with an adjective. The string in (6c) tolerates the emergence of a quantifier, a demonstrative, and an adjective.

Despite these co-occurrences, TsB does not permit demonstratives to coexist with the numeral *jan* or an indefinite quantifier *kra*. Consider (7):

- 7) a. *jan urgaz-ad b. *kra urgaz-ad
one man-this some man-this

The facts in (6) and (7) seem rather complicated under Abney's (1987) analysis where the head D^0 comprises articles, demonstratives, and the genitive. Such a generalization cannot capture TsB data where, for example, noun modifiers show up in a single structure.

In brief, the co-occurrence of nouns and their external modifiers is a highly organized procedure. There are some implicit restrictions on the possible combinations within the nominal structure. Nouns precede adjectives, demonstratives, and clitics. Yet, they follow numerals and quantifiers. These structural arrangements imply that phrasal architecture results from lexical properties of items and syntactic derivational constraints.

I have shown that TsB nouns have nontrivial internal and external structures. The inner makeup consists of the immediate head noun and its integral affixes. The peripheral construction is comprised of the large nominal form as it combines with

its direct outright modifiers. Both of these provisions have close relevance for capturing regularities in formal syntactic derivations.

The Proposal

This section reconsiders the particularities of TsB nominals presented above. It proposes an alternative examination of the noun structure to facilitate the overall characteristics and understanding of TsB nominals. The proposition imposes consideration of the existence of a determiner phrase (DP) in the syntactic derivations of nouns; and advocates that DP is a multi-layered arrangement in TsB. This approach provides a clearer analysis of the specific nature of the phenomenon under scrutiny.

Initial Vowel Status

The status of the initial vowel is, generally, a disputable milestone issue in Berber linguistics. Many researchers consider it to be a determiner category (cf. Ouhalla 1988; Achab 2003). Ouhalla (1988) based his claim on the apparent resemblance between French and Berber determiners. He claims that the two categories equally enclose number and gender information. Likewise, Achab (2003) defends the determiner nature of the initial vowel. He proposes that the vowel heads D^0 . The stem noun is, consequently, its NP complement.

Moreover, Ouhalla (2005) recasts this view and declares that the initial vowel does not correspond to a determiner. He argues that definiteness is a property of discourse rather than morpho-syntax. This assertion is based on the observation that when a noun is assigned a definite reading that interpretation may change due to the presence of other modifiers such as numerals (Compare the structures in (2a) and (5) above.).

In fact, post verbal subjects tend to favor a definite reading interpretation, as in:

- | | |
|---|---|
| 8) a. i-sawl urgaz
he-spoke the-man.CS
“The man spoke.” | b. i-sawl jan urgaz
he-spoke one man.CS
“Some man spoke.” |
|---|---|

In (8a) *urgaz* is interpreted as a definite noun. However, the same form in (8b) is understood as indefinite (Note that *urgaz* in (8b) is also in CS, exactly like the form in (8a). However, it is interpreted as indefinite.).

Within the same spirit, Guerssel (1992, 1995) argues that the first vowel in Berber nominals is a default case marker. This claim develops from his assumption that state forms (FS-CS) are actually case alternations. He, thus, proposes a Kase Phrase (KP) to account for nominal derivations in the language. Under his analysis, D^0 carries gender features. However, when it takes its NP complement in masculine, the initial vowel is associated both with the head D^0 and K^0 , as in:

9) $[_{KP} a- [_{DP} a- [_{NP} -fɾux]]]$

Regardless of the categorial status of this dual head, any raising operation, either to *D* or *K*, generates unattested forms in the system. Likewise, in the feminine, the *t*-heads *D*⁰. Yet, the stem noun is inserted as $[_{NP} -aɾɾux(t)]$. In this instance, the initial vowel *a-* is considered part of the stem noun. This ambivalence remains a puzzling issue under this analysis. On the other hand, El Moujahid (1993, 1997) considers the initial vowel a nominalizer category (cf. Lahrouchi 2013). This view seems to be much closer to the essential specificities of TsB.

Based on the assumptions of El Moujahid (1993, 1997) and Lahrouchi (2013), I consider the initial vowel to be an integral part of the head noun. In this regard, it is also endowed with a default masculine feature interpretation, as has been alluded to in section (“Gender”) (see footnote 1, 3 and 4). One motivation for upholding the substantial nature of noun-initial vowels is the absence of a single example where the initial vowel gets separated from the head noun in TsB. Given this observation, disjunct arrangements such as $[a-[-ɣɟul]]$ “donkey” or $[ta-[-muɾff]]$ “queen / molly” are not attested in the language. This is so, because neither bare *ɣɟul* nor *muɾff* appear in any context in TsB. Yet the whole form *aɣɟul* or *tamuɾff* do exist, and constitute indivisible units, both syntactically and phonologically.

Another argument that supports the intrinsic consistency of initial vowels in nominals comes from vocative constructions. In TsB, a vocative form consists of a vocative particle⁴ *wa-* followed by a noun,⁵ as in:

- | | | | |
|-----|--------------------------|----------------|--------------------------|
| 10) | a. ?a- / wa-(j) *(a)rgaz | b. wa tɪfɾxi:n | c. wa-(j) *(a)ɣɟul |
| | voc. man | voc. girls | voc. donkey |
| | “Hey/O man!” | “Hey/O girls!” | “Hey you stupid person!” |

In (10a), the addressee must appear as *argaz*. The absence of the initial vowel leads to ungrammaticality, as indicated by the asterisk. The same reasoning applies

⁴TsB has other variant vocative forms that are used, mainly, when addressing people older than oneself, as in:

- | | | | |
|----|------------------|-----|-------------------|
| i) | (da)dda ħmad | ii) | taba faɖm |
| | voc.msc. ħmad | | voc.fem. faɖm |
| | “Mr./Dear Ahmad” | | “Ms./Madam Fadma” |

⁵In fact, vocatives may also precede pronouns, as in:

- | | | |
|----|---------------|--------------|
| i) | a. wa-hju | c. wa-hjata |
| | voc. you.msc. | voc.you.fem. |
| | “hey you” | “hey you” |

to (10c). In (10b), the plural form *tifrxɪ:n* cannot be altered, as it is issued from the lexicon. The CF *tfrxi:n* is not allowed in this context.

Since the initial vowel does not adequately stand for a determiner, it is imperative to find out what indicates a *Det* category; especially that a form like *argaz* may mean “the/a man” (Consider examples in (2), (5), and (8a).). To solve this puzzling query, I propose that despite the lack of overt disjunctive articles, there exists a DP projection in TsB (cf. El Moujahid (1993, 1997)). This account explains the ambivalent reading of isolated nouns in TsB.

One argument that supports the presence of an abstract D-category comes from lexical borrowing.

- | | | | |
|-----|-------------|---------|-----------------|
| 11) | a. l-ktab | *ktab | “The Book” |
| | b. l-gamila | *gamila | “Casserole pot” |
| | c. l-mndi:l | *mndil | “Handkerchief” |
| | d. l-makina | *makina | “Machine” |

As is shown by the examples in (11), loanwords are adopted and equipped with their articles (cf. Wexler 1996). The article and the noun, thus, form one integral unit. This connection is real, as the two forms show in other nominal derivations. For example, the plural of *lktab* is *lktub*. **Ktub* without the determiner is unacceptable, as is also indicated by the starred forms in (11). This is evident since the feminine (diminutive) form of *lktab* is *ta-lktab-t* “little book.” Note that feminine affixation applies to the whole form *lktab*, and not simply *ktab*. The same reasoning is accurate with regards to the form *lgamila*. The word is originally feminine. It could be made diminutive by the addition of the feminine morpheme as in *ta-lgamila-t* “little casserole.” The diminutive feminine affixes are attached to the category *lgamila*, not *gamila*. What is of note is when it is made massive: it is realized as *a-lgamil* “large casserole.” The same process is again applicable. Notice that the masculine (massive) *a-* is attached to the word *lgamil*, which already contains a definite article, as indicated in (11b).

As is evident from this reflection, although borrowed words are viewed as enclosing two distinct components in the source language, they are considered as one complete entity in TsB. Any attempt to disentangle the determiner and the noun leads to ungrammaticality.

In addition, all of these, and similar, words are definite in TsB even when they appear in isolation. However, they can be made indefinite in presence of numerals, for example, consider (12):

- | | | |
|-----|----------------|-----------------------|
| 12) | a. Ja(n) lktab | b. jat lgamila |
| | One the-book | one the-casserole pot |
| | “One book” | “One casserole pot” |

This observation further supports the claim that all nouns are viewed as one single integral unit. The language does not allow isolated articles. However, this apparent determiner-less characteristic does not preclude the projection of a D-category in the

syntax of the language. This view has positive advantages leading to the discovery of a number of properties about the syntax of nominals in TsB.

The DP Organization in TsB

Given the particular facts examined above, I suggest that TsB nouns have one single unified and layered syntactic structure, as in:

$$(13) \quad [_{DP} D^0 [_{nP} N^0 -n^0 [_{NP} \mathfrak{N}^0]]]$$

The proposal in (13) is motivated by the dominating assumptions in MP, whereby derivations reduce to the essential process that is needed for explaining the conceptual-intentional machinery of humans. Therefore, TsB noun representations consist of a basic structure that is exemplified in (13). In this regard, and as is presumed in the relevant literature, I assume that all NPs merge with a functional category n (This is an abstract form. It is similar to little v in vP -VP analyses.) where the former function as complements of the latter. This is so because n has a selectional feature $[u-N]$, which ensures merger with NP. The formed nP is then amalgamated with a head D , and a DP is projected.

This kind of representation evokes the underlying economy principles of MP, which require reduction of operations and representations except for indispensable material (cf Chomsky 1995). To eliminate redundancy, I associate n with a valued copy of ϕ -features (cf. Pesetsky and Torrego 2007). This approach asserts avoidance of superfluous syntactic representations, whose functions can simply be obtained otherwise, without recourse to projecting multi-layered syntactic nodes (It is tempting to have a gender Phrase (GenP) and a Number Phrase (NumP) project above NP (cf. Ritter 1991). However, I consider them to simply be idiosyncratic properties of nominals, as they are parts of their lexical entries. Hence, gender and number, together with the categorial feature $[N]$, are licensing conditions for nouns, which are introduced through n^0 , in syntax.).

Actually, this analysis of n portraying nominal features has first been implemented in Lowenstamm (2008). On the basis of Distributed Morphology approach, he claims that n is headed by gender features of the noun in French, as well as in Yiddish. Lahrouchi (2013) goes a step further in analyzing nouns in TsB and claims that little n incorporates both gender and number features. In terms of templatic morphology, he demonstrates that gender and number affixes are correlated with a prefixal prosodic pattern consisting of a CV string (cf. Lowenstamm 2008). The mapping to syntax associates the CV structure with n^0 head of nP . Inspired by these analyses, I suggest that n acts as a licenser for the noun element. In view of that, and in addition to its categorial feature $[N]$, little n is endowed with ϕ -features. The head n is, thus, a nominalizer that licenses the NP complement, in terms of s-selection. I endorse this approach, as it establishes a satisfactory account of nominal representations in TsB syntax.

N⁰-to-n⁰ Raising

Under (13), the head noun is specified for an $[i-N]$ feature. Yet, n has an uninterpretable copy of that feature, $[u-N]$. This fact forces N to adjoin to n overtly. This process is obligatory as n needs to get rid of the non-interpretable N feature (cf. Wurmbrand 2014; Pesetsky and Torrego 2006). For that reason, head nouns are forced to move to n in order to match and eliminate the uninterpretable categorical and phi-feature: $[u-N]$ $[u-\phi]$ (I believe that Once N targets n , the merging process results in an extension of the root node NP. Therefore, all the nodes above NP are a simple expansion of the primary source node. This is true, because the ultimate procedure of phrase derivation is always interpreted to be equivalent to the initial basic form. Hence, a nominal form will constantly end up being interpreted as an enlargement of a noun. This judgment circumvents a variety of restrictions on head movement constraints, recently addressed in the literature concerned.). This procedure is, empirically, in accordance with the facts of the language. Consider the following structures:

- | | | | |
|--------|---|----|---|
| 14) a. | argaz igzzuln
man.sg short.sg
“A/the short man” | b. | tamyart ihlan
woman.sg beautiful.sg
“A/the beautiful woman” |
| c. | irgazn gzzulnin
men short.pl
“(the) short men” | d. | tumyarin hlanin
women beautiful.pl
“(the) beautiful women” |

The forms in (14), as has already been shown in (3) and (6) above, show that adjectives follow head nouns in TsB. This order is fixed in the language. Thus, an *Adjective-Noun* order is unacceptable. This observation supports the claim that N adjoins to n in TsB, as (13) instantiates.

Accordingly, the structures in (14) are generated by projecting and NP headed by a noun. Following the influential work of Cinque (1994), I consider TsB adjectives to be merged as specifiers of NPs (They might also be simply attached as adjuncts to the noun. Although this issue is controversial in the literature, it does not affect the line of reasoning in this work.). This left attachment imposes that the post-nominal occurrence of adjectives in (14) does involve N-raising (It seems that N-movement within a DP occurs in analogy with verb movement in a clause. The speculation is that languages which necessitate verb movement also force N-raising to take place. On the other hand, languages without overt V-movement lack N-movement as well. A good example could be English and TsB. In English, the verb does not rise to T , overtly. Therefore, adjectives appear before nouns. In TsB, sentential word order is dominantly VSO which indicates overt V-raising to T . For that reason, adjectives are always post-nominal.) to a position higher than NP, as represented in (15).

- | | | |
|-----|--|--|
| 15) | $[_{NP} N^0 - n^0$
$[\#N][\# \emptyset]$
$[val]$ | $[_{NP} AdjP \text{ } N^0]]$
$[iN] [i-\phi]$
$u[val]$ |
|-----|--|--|

Under this proposal, *N* is drawn from the lexicon, fully specified with its syntactically relevant features. However, the latter need to be licensed according to the derivational procedures. Incidentally, N^0 is enumerated with unvalued interpretable categorial (*N*) and phi-features (φ). Little *n* is also specified for identical, yet uninterpretable and valued, features. The valued phi-features of *n* certify the activation of the corresponding features on the head noun. Since the theory requires irrelevant material to be eliminated at the interfaces, the uninterpretable features of *n* have to be checked and removed. For that reason, *N* raises to *n*, leaving the adjective within NP. According to this process, the non-interpretable [*N*] and [φ] features of *n* are matched, and deleted. This course of action results in the acceptable order in (14).

However, the agreement on the adjective remains unexplained (It may be possible that subsequent to merger, the head noun, and the adjective in the NP specifier, trigger a spec-head agreement prior to *N* raising to *n*). As is clear from (14), TsB adjectives show partial concord with the nouns they modify. They only agree in number. Therefore, adjectives have an [*uNum*] which needs to match the one on *N*. Once *N* adjoins to *n* to check/valuate its categorial and phi-features, it c-commands the adjective. In this configuration, it spreads its valued [*iNum*] to the adjective. The two [*Num*] features on *N* and the adjective get matched and valued in this domain. This process results in the sharing of number features between the two categories.

Given the provided data, the ordering of post-nominal adjectives is most excellently accounted for by postulating an *nP* above NP. The specifier of the latter encompasses adjectives. *N*-raising to *n* generates the right word order.

N^0 - (*n*)-to- D^0 Raising

On the basis of (13), there exists a DP projection in TsB. Indeed, the facts of the language suggest that TsB recognizes the presence of a null *D*-category at the interpretive interface. This is so because *D* contains (in)definiteness features [$\pm\text{Def}$] (The grammatical concept of (in)definiteness seems, thus, to be universal regardless of morphological realization. It is, consequently, omnipresent in natural languages regardless of overt spell out.) (cf. Fassi Fehri 1999 and Benmamoun 2000). Since this feature has semantic content, it is interpretable on D^0 . Hence, the fact that TsB lacks plain articles does not confirm the absence of a DP projection.

In conformity with the [$\pm\text{Def}$] specification, I propose that there are four categories that head D^0 : the demonstratives *-ad* “this” and *-ann* “that,” the quantifier *Kra* “some,” the numeral *jan* “one,” or a phonetically void element \emptyset .

The Demonstrative

Demonstratives (*Dem*) are deictic categories in TsB. Regarding morphology, they are an invariable class. (They are also affixal, as they cannot stand alone. They are required to suffix to a head noun.) They do not show agreement with the head noun to which they attach. They basically express proximity from the speaker.

- | | | | | |
|-----|-------------|--------------|--------------|---------------|
| 16) | a. argaz-ad | b. irgazn-ad | c. argaz-ann | d. irgazn-ann |
| | Man-this | men-this | man-that | men-that |
| | “This man” | “These men” | “That man” | “Those men” |

As is clear from (16), demonstratives are post-nominal categories. Moreover, they express definiteness. All the nouns in (16) are understood to be definite (Consider the examples in (7a–7b).). The reason is that the $[\pm\text{Def}]$ feature is interpretable on the noun. In other words, it is part of its lexical entry. However, that very feature may not be valued from the lexicon. This is the reason why a noun like *argaz* in isolation may be interpreted, either as “*a man*,” or “*the man*.”

Given this state of affairs, let us consider the derivation of the nouns in (16). First, they merge with little *n*. The latter has non-interpretable categorial and phi features that need to be checked and deleted. The head noun is then forced to adjoin to *n*, as in (15) above. The *nP* projection merges with D^0 headed by the demonstrative due to interpretable (in)definiteness feature on N. The head of DP (*–ad* or *–ann*) is specified with a non-interpretable $[\mu\text{Def}]$ feature and an unvalued $[\text{Def}]$ feature. The noun, which has adjoined to *n*, has an interpretable $[i\text{Def}]$ feature and a valued $[\text{Def}]$ one. The complex head *N–n* raises to D. In doing so, it values the unvalued $[\text{Def}]$ feature and eliminates the non-interpretable $[\mu\text{Def}]$ feature of D.

One empirical argument that supports this analysis comes from forms like (6b–6c), repeated here as (17a–17b)

- | | | |
|-----|---------------------|-----------------------------|
| 17) | a. argaz-ad igzzuln | b. kullu irgazn-ad gzzulnin |
| | man-this short.Sg. | all men-these short.PL |
| | “This short man” | “All these short men” |

The occurrence of *N–Dem* before the adjectives indicates that the head noun underwent two movement operations. The first one is from within NP into *nP*: this adjunction results in leaving the adjective behind within NP. The second is that the complex *N–n* has adjoined to the demonstrative in D^0 . The same process functions for (17b). The appearance of the quantifier, *kullu*, indicates that it is merged in spec DP. (See footnote 6 below.)

The Quantifier

Another element that occupies D^0 is the word *kra* “some.” Unlike the demonstratives, this item can stand alone. In accordance with the grammatical attributes of TsB, this word is an indefinite quantifier. Semantically, it is a nonspecific item. Therefore, the nominal form that follows it is constantly interpreted as indefinite and unspecific, as is clear from (5b), repeated in (18a).

- | | | |
|-----|--------------|------------------|
| 18) | a. kra urgaz | b. *kra urgaz-ad |
| | some man | some man-this |
| | “Some man” | *“This some man” |

The nominal form *argaz* in (18a) is an unidentifiable person. It can be equated with the indefinite pronoun *somebody*. The reason behind this statement is that *kra* is introduced in D^0 and it is specified with a non-interpretable and unvalued [Def] feature. The head noun has an interpretable [Def] which is valued. Therefore, the uninterpretable feature needs to be removed; and the unvalued one needs to be valued prior to transfer.

So, when *N* adjoins to *n*, the *nP* projects and merges with D^0 , which is headed by *kra*. Since the latter is not affixal, it blocks the complex *N-n* to raise to *D*. This blockage is justified by the ungrammaticality of **urgaz kra*, where *urgaz* moves to *D*. Yet, the predicament of features needs to be explained, as (18a) is grammatical.

Since the complex *N-n* cannot raise to *D*, an AGREE relation between the latter and the former is established. The syntactic computation proceeds as follows: the head *kra* probes down for the goal *urgaz*⁶ within *nP*. The latter is c-commanded by the former, hence, the features of *D* are matched with those of *N*. This process hands the derivation to the interfaces.

One argument that supports this conception comes from the ungrammaticality of (18b). The form is unacceptable because of (in)definiteness feature mismatch between the demonstrative *-ad* and the indefinite quantifier *kra*: the former is [+Def], the latter is [-Def]. The structure, thus, contains two distinct values of (in) definiteness feature. Since the nature of features is crucial in derivations, the [Def] divergence in (18b) results in ungrammaticality.

⁶It is worth pointing out that the variation between *argaz* and *urgaz* is a state alternation. *argaz* is the free state form, while *urgaz* appears in the construct state. I believe that construct state forms occur with the presence of a genitive preposition *n*. the latter is endowed with a peculiar lack of restrictions. Sometimes, it is overt; other times it is covert.

- | | | | | |
|----|-------------------|---------------------|--------------|-----------------|
| 1) | a. tiggmi n-urgaz | b. tiggmi n-tmyart | c. kra urgaz | d. kra n-tmyart |
| | house of-the-man | house of-the-woman | some man | some of-a-woman |
| | "the man's house" | "the woman's house" | "some man" | "some woman" |

In genitive forms (1a–1b), the preposition is obligatory. In quantified forms (1c–1d), *n* is obligatory when the following singular noun begins with a consonant, as in (1d); or *kra n lxñšt* "some bag." It is also obligatory with plural nouns, as in: *kra n ırgazn* (realized phonologically as *kra ıı ırgazn*, "some men"). The string in (1c) is thus best represented as *kra (n) urgaz*, where the preposition *n* is not phonetically realized. In this regard, the preposition *n* heads a *KP*, as it is a Case assigner category. It, thus, takes *DP* as its complement, as in:

- 2) [_{KP} n [_{DP} jan/kra [_{nP} N-n [_{NP} N]]]]

In order to obtain the correct word order, the whole *DP* in (2) adjoins to *KP*. This process ensures obtaining (1d), for example. Due to lack of space, I leave this issue to a future exploration.

The Numeral

Another category that occupies D^0 is the term *jan*. It is a numeral category that serves to provide information about the complete number of countable nouns, as in (5a), repeated in (19a).

- 19) a. *jan urgaz* b. *jat tmyart* c. *sin irgazn* d. *snat tmyarin*
 one.ms. man one.fm.woman two.ms men two.fm. women
 ‘‘One man’’ ‘‘one woman’’ ‘‘two men’’ ‘‘two women’’

As is clear from (19), numerals vary in gender specification. They agree with the following noun in its gender feature. The legibility implied in this observation is that the semantically interpreted gender on the noun is reflected on the morphological specification of the numeral. Any feature discrepancy is thus expected to trigger ungrammaticality, as in:

- 20) a. **jat argaz* b. **jan tmyart* c. **sin tmyarin*
 one.fm. man one.ms.woman two.ms women

The explicitness evoked in (20) is that nominal features, and determiner features, need to line up compatibly in the strings where they appear; otherwise, feature value dissimilarity ends up in ungrammaticality.

Methodically, the derivations of the forms in (19) set off by projecting an NP, headed by an N category. N is specified with its categorial and phi-features. The projection merges with the head *n* identified with non-interpretable categorial and phi-features. This condition forces N to internally merge with *n*. The process winds up in the elimination of uninterpretable features. The *n*P projects and it merges with D^0 (numeral), due to the presence of (in)definiteness features on N(-*n*).

One argument that sustains this view emanates from the following constructions:

- 21) a. *jan urgaz iħrurin* b. *jat tmyart iħrurin*
 One.ms man beautiful one.fm. woman beautiful
 ‘‘One handsome man’’ ‘‘One beautiful woman’’

The occurrence of the noun in pre-adjectival position indicates that it has moved to adjoin to *n* within *n*P. Moreover, its presence in post numeral site suggests that it did not move into D. the reason for that blockage is due to the non-affixal nature of the numeral and its occupying the head D. This idea is justified by the ungrammaticality of (7a), repeated below as (22).

- 22) **jan urgaz-ad*
 One.ms. man-this

The ungrammaticality of (22) stems from a feature mismatch between *jan*, which is [−Def] and *−ad* which is [+Def]. A form cannot be definite and indefinite at the same time. Likewise, (22) is unacceptable by reason of the confusion it engenders.

The construction contains two D categories: *jan* and *-ad*, which is illegitimate. The appearance of the noun before the demonstrative suggests that the former has moved to D. In the same way, the manifestation of the numeral in the beginning of the string indicates that it occupies D, and that N cannot move into D. This confusion generates lack of certainty about the makeup of the lined up words. Thus, (22) is ruled out as ungrammatical.

Given this conception, the irrelevant indefinite feature of D in (19) and (21) are eliminated through an AGREE relation between D and N. This reasonable analysis is attributable to the illegitimacy of N-*n* raising. The non-interpretable and unvalued [*-Def*] features of *jan/jat*, force D to probe down for the goal *urgaz/tmyart* within *nP*. AGREE operates, matching the feature of the noun and the determiner; and the impertinent features get deleted accordingly. The formed DP is, subsequently, transmitted to the interfaces.

Numerals are indefinite determiners. Since they are non-affixal entities, they block N-raising to D. The uninterpretable and unvalued features of the latter are checked/valued, and deleted through the concept of AGREE.

The Null Category

In conformity with the DP hypothesis, NPs are generally complements of D (In the analysis presented in this chapter, *nP* is the actual complement of D. However, I consider *nP*, as well as DP, as extensions of NP. Therefore, NP could always be regarded as the complement of D.). Since article-less languages, such as TɔB, exist, it is possible that DP can be headed by a phonologically empty head. One reason for this claim could be that D subcategorizes for NP. This implies that the presence of N entails the existence of a c-commanding D⁰ above it. If this claim is true, all bare nouns have a DP structure, as presented in (13) above.

To verify the prospect of this supposition, let us consider (2a), above, repeated here as (23).

- 23) i-mmudd-a urgaz
 he-travel-perf. the-man.CS
 ‘The man travelled.’

The subject (*urgaz*) in this sentence is a bare noun, as it surfaces without any external modifiers. The first interesting observation is that it is interpreted as definite category. If it is assigned an indefinite reading, the sentence becomes ungrammatical. This recognition, therefore, necessitates that the head noun bears [*+Def*] feature. This property can only be factual if N has moved to D, as discussed above.

The derivation of the subject in (23) proceeds on the basis of the representation in (13). The head noun projects an NP structure. The latter merges with little *n*. The head *urgaz* is obliged to adjoin to *n* to check/value the nominal categorial feature and eliminate the non-interpretable phi- features on *n*. Since N is specified with definiteness features, it moves to D, and its equivalent, uninterpretable [*Def*] feature, gets deleted.

This view is empirically supported by sentences where adjectives or demonstratives occur.

- 24) a. i-mmudd-a urgaz igzzuln
 he-travel-perf. the-man.CS short
 “The short man has travelled.”
 b. i-mmudd-a urgaz-ad igzzuln
 he-travel-perf. man-this.CS short
 “This short man has travelled.”

The presence of the adjective in a post nominal position in (24a) indicates that *urgaz* has moved out of NP into, at least, *nP*. The adjective is left within NP. As has been argued above, the demonstrative *-ad*, in (24b), merges as the head of D. The appearance of *urgaz* to the left of it (*urgaz-ad*) suggests that the former has adjoined to D, overtly.

This proposal is further supported by the presence of a possessive adjective in a nominal structure like (25).

- 25) i-mmudd-a urgaz-n-iw
 he-travel-perf. the-man.CS-of-me
 “My husband has travelled.”

In their most essential aspect, possessive adjectives, like *n-iw* in (25), are a shortened form of a possessive preposition *n-* “of,” and its nominal object. Let us exemplify this idea in (26):

- 26) a. zri-γ argaz n-xlizza b. zri-γ argaz-n-ns
 Saw-I the-man of-Khadija saw-I the-man-of-her
 “I met the husband of Khadija.” “I met her husband.”

As is evident from (26), *n-ns* in (26b) replaces *n-xlizza* in (26a), especially that both forms are preceded by the same category. The noticeable difference is that *n-xlizza* is an independent construction, while *n-ns* is not. Therefore, the possessive adjective is an enclitic form. Being so, it cannot stand alone. As a consequence, it needs to be hosted by the preceding category, at least at the PF interface.

Since both *n-xlizza* and *-n-ns* are posterior to the object *argaz* in (26), I assume that they are merged into the same position. They, equally, function as complements of the object noun, as sketched in:

- 27) [_{DP} D⁰ [_{nP} n⁰ [_{NP} N⁰ [_{PP} *n-xlizza* /-ns]]]]

This representation is precisely the same schemata form in (13), supplied with a prepositional phrase as a complement of the noun. However, a little distinction needs to be pointed out here. This concerns the fact that *n-xlizza* in (26a) stays in its initial merging position. This claim is supported by word order, as in: *argaz-ad igzzuln*

n-xlizza. The PP is preceded by the adjective *igzzuln* which occupies spec NP, and the noun-demonstrative amalgam which heads D^0 . Yet, the situation with the enclitic *-n-ns* is different, as (28) illustrates.

- 28) a. *zri-γ argaz-n-ns igzzuln* b. *zri-γ argaz-ad-n-ns igzzuln*
 saw-I the-man-of-her short saw-I man-this-of-her short
 “I saw the short husband of hers.” “I saw this short husband of hers.”

Comparing (28a) and (28b) shows that *-n-ns* raises upwards, and adjoins to the object *argaz*. Yet, the nature of this adjunction is somehow puzzling, as it involves syntactic, as well as post-syntactic derivations. In syntax, *-n-ns* merges as the complement of *argaz*. However, due to its being a dependent and affixal element, it requires a host. After N raising to *n*, and, subsequently, to D, the potential host is lost.

I propose that *-n-ns* adjoins to NP, thus, making (27) look like (29).

$$(29) \quad [_{DP} D^0 [_{nP} n^0 [_{NP} -n-ns [NP N^0 [_{PP} -n-ns]]]]]$$

It seems highly desirable to maintain the representation in (29).⁷ The PP *-n-ns* is a maximal projection allowed, as such, to only conjoin to top nodes, as the adjunction condition requires. The adequately available spot is the NP layer, as is indicated in (29). Adjunction to this position results in ordering the enclitic, *-n-ns*, anterior to the adjective within the lower NP and posterior to the noun in D, as (28) shows. However, *-n-ns* cannot immediately encliticize into the preceding noun overtly. This is so because the *nP* phrase intervenes between the head noun in D, and the

⁷It is appealing to recommend that the enclitic element is better if it ends up in the available spec *nP*, thus every part of the derivation takes place without recourse to PF operations, as in:

$$i) \quad [_{DP} D^0 [_{nP} -n-ns [_{nP} n^0 [_{NP} N^0 [_{PP} -n-ns]]]]]$$

However, this proposal faces a particular number of problems. One of them is the impossibility of deriving:

- ii) *Kra ufrux-n-ns*
 Some son-of-hers
 “One of her sons”

The derivation is going to always generate ungrammatical forms such as:

- iii) **Kra -n-ns ufrux*
 Some -of-hers son

Forms such as (iii) are unattested in the language.

adjective with NP. It is only after spell-out, at the PF interface, where the hierarchy of syntax is no longer required, that the two forms get combined. Parsing at the PF detects the abandoned enclitic form unhosted, and forces incorporation into the closest preceding host.

As has been shown, it is possible that DP projects while its head may not be occupied by discernible categories. Although D may remain phonetically empty, it must be syntactically legitimized. It can only be so by the interpretable features it hosts, namely, (in)definiteness specifications.

On the whole, and in conformity with the DP hypothesis, I have demonstrated that DP projects in TsB. Its head bears (in)definiteness that legitimizes its licensing. I have shown that when D is definite, it either hosts an affixal demonstrative (–ad), or remains phonetically void, yet interpretably active with a [+Def] feature. When it is indefinite, it is occupied, unavoidably, by an indefinite quantifier or a numeral; otherwise, it is enriched by the semantically active [–Def]. These conditions rule out the option of absence of DPs from the inventory of article-less languages, such as TsB.

Conclusion

All things considered, the chapter has particularly presented a new insight into the structure of nominal forms in TsB. It has tackled the internal and external organization of the category under analysis. The inner and outer makeup are interconnected, and favor reduced projections in the functional domain of the noun. This perspective challenges most of the proposals in the literature. The chapter has shown that the initial vowel is an integral masculine part of the noun. Moreover, it has established that phi-features do not necessarily correspond to prolifically iterated projections. I have argued that DP projects in the language regardless of the absence of articles in the system. The presence of a D⁰ is motivated by (in)definiteness features, which grant reference to nouns. In doing so, I have provided an innovative explanation for the syntax of pre- and post-nominal modifiers in the language.

References

- Abney S (1987) The English noun phrase in its sentential aspects. Ph.D. dissertation, MIT
- Achab K (2003) Alternation of state in Berber. In: Lecarme J (ed) *Research in Afroasiatic Grammar II: selected papers from the fifth conference on Afroasiatic Languages*, Paris 2000. John Benjamins, Amsterdam, pp 1–19
- Alderete J, Jebbour A, Kachoub B, Wilbee H (2015) Tashlhiyt Berber grammar synopsis. http://www.sfu.ca/~alderete/datasets/aldereteEtal2015_tashlhiytGrammarSynopsis.pdf. Accessed July 2020
- Arkadiev P (2015) The Berber “state” distinction: Dependent marking after all? A commentary on Mettouchi and Frajzyngier (2013). *Linguist Typol* 19(1):87–100
- Bendjaballah S, Haiden M (2005) *The Grammar of prepositions in Berber*. (Taqbaylit). Ms. CNRS et Université Lille 3
- Benmamoun E (2000) *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford University Press, Oxford

- Bensoukas K (2018) Concurrent cognate and contact-induced plural traits in Afro-Asiatic: Amazigh *id-* and Arabic *-at* Plurals. *IJAL* 4(1):59–102
- Bošković Z (2005) Left branch extraction, structure of NP, and scrambling. In: Sabel J, Saito M (eds) *The free word order phenomenon: its syntactic sources and diversity*. Mouton, Berlin, pp 13–73
- Bošković Z (2008) What will you have, DP or NP? In: Elfner E, Walkow M (eds) *Proceedings of the Northeastern Linguistics Society (NELS) 37*. GLSA Publications, Amherst, pp 101–114
- Boukhris F, Boumalk A, Moujahid E, El-H., and Souifi, H. (2008) *La nouvelle grammaire de l'Amazighe*. Publications de l'Institut Royale de la Culture Amazighe. El Maarif Al Jadida, Rabat
- Chomsky N (1995) *The minimalist program*. MIT Press, Cambridge, MA
- Chomsky N (2001) Derivation by phase. In: Kenstowicz M (ed) *Ken Hale: a life in language*. MIT Press, Cambridge, MA, pp 1–52
- Cinque G (1994) On the evidence for partial N-movement in the Romance DP. In: Cinque G, Koster J, Pollock J-Y, Rizzi L, Zanuttini R (eds) *Paths towards universal grammar: study in honor of Richard S. Kayne*. Georgetown University Press, Washington, pp 85–110
- El Moujahid El-H (1993) *Syntaxe du groupe nominal en berbère Tachelhiyt (Parler d'Igherm, Souss, Maroc)*, Doctoral thesis, Mohammed V University, Rabat
- El Moujahid El-H (1997) *Grammaire générative du berbère: Morphologie et Syntaxe du Nom en Tashelhit*. Rabat and Casablanca: FLSH [Série Thèses et mémoires 38]
- Ennaji M (2001) The construct state in Berber. *Stud Linguist Sci* 31(2):55–72
- Fassi Fehri A (1993) *Issues in the structure of Arabic clauses and words*. Kluwer, Dordrecht
- Fassi Fehri A (1999) Arabic modifying adjectives and DP structures. *Studia Linguistica* 53(2): 105–154
- Felice L (2020) On the case system of Kabyle. *McGill Working Papers Linguist* 26(1):1–21
- Guerssel M (1983) A phonological analysis of the construct state in Berber. *Linguist Anal* 11(3): 309–329
- Guerssel M (1992) On the Case System of Berber. *Can J Linguist* 37(2):175–195
- Guerssel M (1995) Berber Clitic doubling and syntactic extraction. *Revue québécoise de linguistique* 24(1):111–133
- Idrissi A (2000) On Berber plurals. *Research in Afroasiatic Grammar*. In: Lecarme J (ed) *Papers from the third conference on Afroasiatic languages*. John Benjamins, Amsterdam/Philadelphia, pp 101–124
- Kossmann MG (2009) Loanwords in Tarifit, a Berber language of Morocco. In: Haspelmath M, Tadmor U (eds) *Loanwords in the world's languages. A comparative handbook*. Mouton, Berlin, pp 191–214
- Kossmann MG (2014) Derivational gender in Moroccan Berber: examples from Ayt Seghrushen. *STUF-Lang Typol Univers* 67(1):21–33
- Lahrouchi M (2013) Templates, markers and syntactic structure in Tashlhiyt Berber. *Lingua* 133: 53–72
- Longobardi G (1994) Reference and proper names: a theory of N-movement in syntax and logical form. *Linguist Inquiry* 25(4):609–665
- Lowenstamm J (2008) On little *n*, $\sqrt{\text{ }}$, and Types of Nouns. In: Hartmann J, Hegedus V, van Riemsdijk HC (eds) *Sounds of silence: empty elements in syntax and phonology*. Elsevier, Amsterdam, pp 106–143
- Lyons C (1999) *Definiteness*. Cambridge University Press, Cambridge
- Mourigh K (2015) *A grammar of Ghomara Berber*. PhD thesis. Leiden University, Leiden
- Nejme FZ, Boulaknadel S, Jdine DA (2013) Finite state morphology for Amazighe language. In: *Proceedings of the international conference on intelligent text processing and computational linguistics 1*, Lecture notes on Computer Science book series (LNTCS, volume 7816). CICLing, Samos, pp 189–200
- Ouhalla J (1988) *The syntax of head movement: a study of Berber*. Ph.D. Thesis, UCL

- Ouhalla J (1996) The construct state in Berber. In: Lecarme J, Shlonsky U (eds) *Studies in Afroasiatic Grammar*. Holland Academic Graphics, The Hague, pp 278–301
- Progovac L (1998) Determiner phrase in a language without determiners. *J Linguist* 34:165–179
- Pesetsky D, Torrego E (2006) Probes, goals and syntactic categories. In: Otsu Y (ed) *Proceedings of the 7th Tokyo conference in psycholinguistics*. N.P, Tokyo, pp 25–60
- Pesetsky D, Torrego E (2007) The syntax of valuation and the interpretability of features. In: Karimi S, Samiian V, Wilkins W (eds) *Phrasal and clausal architecture: syntactic derivation and interpretation*. John Benjamins, Amsterdam, pp 262–294
- Ritter E (1991) Two functional categories in noun phrases: evidence from modern Hebrew. *Syntax Semantics* 25:37–62
- Taghbalout I, Ataa Allah F, El Marraki M (2015) Amazigh noun inflection in the universal networking language. *Int J Educ Inf Technol* 9:122–128
- van Putten M (2016) The origin of front vowel nominal prefixes in Berber. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 116(1):11–39
- Wexler P (1996) *The Non-Jewish origins of the Sephardic Jews*. University of New York Press, New York
- Wurmbrand S (2014) The merge condition: A syntactic approach to selection. In: Kosta P, Franks SL, Radeva-Bork T, Schürcks L (eds) *Minimalism and beyond: radicalizing the interface*. John Benjamins, Amsterdam, pp 130–166

Two Waves of Berber Influence on Moroccan Arabic

24

Jeffrey Heath

Contents

Introduction	518
In the Beginning	518
Thereafter	521
Hereafter	522
References	522

Abstract

The influence of Berber syllabification and phonotactics on Maghrebi Arabic is palpable. However, its influence on Moroccan Arabic (MA) specifically is best divided into two stages: (1) during the brief formative period of archaic MA, and (2) its spread over the succeeding centuries – and still ongoing. In the first stage, the key event was the local interaction of Arabized Berber troops with speakers of Late Latin in the Roman triangle defined by Tangier, Salé, and Volubilis. This combination produced a vocalic system of three full vowels {u a i}, phonemic stress, schwa in unstressed final syllables, limited syncope, and absence of vowel length, which are still observable in northern Morocco. In the second stage, initially in Algeria, and then spread into Morocco, Berber influence led to the vocalic system of the familiar MA koiné: three full (or long) vowels, short vowel(s) schwa, and marginally short ŭ, expanded syncope and absence of stress.

Keywords

Moroccan Arabic · Berber · Vocalic systems

J. Heath (✉)

University of Michigan, Ann Arbor, MI, USA

e-mail: jheath@umich.edu

Introduction

The issue of Berber prosodic and phonotactic influence on Moroccan Arabic (MA) cannot be studied apart from the historical dialectology of MA itself. Unfortunately, almost every report on MA generally or on specific MA dialects has misrepresented the vowel system. This comment applies even to centrally important scientific works on MA, from W. Marçais' ground-breaking study of Tangier (1911) through Harrell's works, including his dictionary (1966), and on to the present day. Indeed, most of these writers acknowledged in print their confusion about the vowel system (see Heath 2018 for references and discussion, which are not repeated here).

In this chapter I attempt to clarify the situation by distinguishing the role of Berber in the brief formative period of archaic MA, chiefly in and just before the eighth century, from its role in the long-term evolution of MA leading to the modern koiné.

In the Beginning

The Arab occupation of Morocco took a very different form from the Arab expansion into other areas. In Egypt and the Levant, the actual Arab military conquest was preceded by a long period of trade and diplomatic contacts, constituting networks that could be built on. In the central Maghreb, a new Arab city (Kairouan) was built in order to remain independent of existing polities (Byzantines, Berber chieftains). Initially for Kairouan, as previously for the Romans and Byzantines. Morocco was a peripheral outpost of no great importance except as a military flank to be defended. Morocco was not actually occupied (by an Arab-led but mainly Berber army) until the last 2 years of the seventh century. Of the small number of ethnic Arabs who remained there after the occupation, many joined the invasion of Andalus, a much richer prize, a little over a decade later.

Instead of immediately building a new Arab city, the occupying troops settled down in the existing Roman cities and towns, chiefly Tangier, and its satellites, plus Salé on the Atlantic coast, and Volubilis (now a ruin near Meknes) in the interior. Fes was not even founded until 789, and not significantly populated until 817–824 with the first wave of immigration from Andalus and Kairouan. There is general agreement among archeologists and epigraphers that Christianity was still practiced, and Late Latin was still spoken, by Christian Romans in these communities at the time of the Arab occupation. There is also agreement among historians that the bulk of the Arab army consisted of Berber troops, initially recruited from the central Maghreb, and later, more locally. These Berbers were partially Arabized linguistically, using a contact Arabic as the language of command and as a *lingua franca*.

These Arabized Berbers cohabited with Late Latin-speaking Roman women, and their offspring became the first native speakers of archaic MA. The women must have played an especially important role in the linguistic development of the first few generations of MA speakers given the high turnover rate among soldiers (invasion of Andalus 710–711, devastating Kharijite revolt 740).

All of this is common sense given what archeologists and historians have told us. Remarkably, the important contribution of Late Latin in the formation of MA has been almost completely overlooked by linguists, who have nevertheless been unanimous in recognizing a powerful Berber influence on the language. Indeed, a degree of cultural antipathy against recognizing a European, as opposed to Afroasiatic, influence on MA has led to more than a century of vigorous denial that MA D-possessives have a Latinate source, offering various far-fetched comparisons to Aramaic or South Arabian forms with no explanation as to how they bypassed the rest of North Africa on the way to archaic MA and early dialectal Spanish Arabic and from there onto scattered pockets of Andalusí commercial activity in coastal and subcoastal Algeria. Yet *d* or archaic *di* before nouns is a perfect formal and functional match for Latin *dē*, reflected throughout western Romance (e.g., Spanish *de*) as the basic possessive linker. Likewise, the allomorph *dyal-*, used chiefly before pronominal suffixes (archaic MA *dyal-u* “of him” and *dyal-a* “of her”) is easily explained as the combination of Latin *dē* plus third person pronouns *ille*, *illud*, and *illā*, subsequently spread, by analogy, to first and second persons (Heath 2015).

Recognition of Late Latin influence on MA phonology and prosody has been delayed by another type of cultural antipathy, this time expressed as a refusal by Marçais and a century’s worth of successors (each predisposing the next) to believe that northern MA dialects have a vowel system structurally distinct from that of the koiné, and of classical and literary Arabic.

We can make only educated guesses about the vowel systems of seventh–eighth century Berber languages, Late Latin, and Berber-influenced contact Arabic. It is generally agreed that Berber languages of the region were less advanced than now in the reduction of short vowels to schwa or (via syncope) to zero; southern Berber (Tuareg) still distinguishes short *ā* from *ə*. So a reasonable starting point is a system like (1):

- (1) Berber vocalism around 700
 - *{u a i} full (long) vowels
 - *{ā ə} short vowels
 - phonemic stress absent

The Arabic spoken by native Arabs of the Kairouan area at that time can be modeled as (2), with the proviso that short-vowel contrasts were likely already becoming blurry. Specifically, the opposition **ā/ī* may already have been neutralized as **ə*, and **ū* may have begun losing ground.

- (2) Arabic vocalism around 700
 - *{ū ā ī} long vowels
 - *{āy āw} diphthongs
 - *{ū ā ī} short vowels, tending toward *{ū ə}
 - phonemic stress absent

This author has argued that Arabized Berbers, as troops in Arab-led armies of the period, spoke a contact Arabic with the features in (3).

- (3) Berberized contact Arabic around 700
 *{ū ā ī} long vowels
 *{āy āw} diphthongs, of variable pronunciation
 *{ū ə} short vowels
 phonemic stress absent

There are of course no documents in Late Latin, from Morocco or anywhere else at this period. Based on early Medieval Spanish and Portuguese, I suggest (4). It's an open question whether there were five vowel qualities (as in modern Spanish) or seven (as in modern Portuguese and Italian).

- (4) Late Latin around 700
 *{u o ɔ a e i} or *{u o a e i} vowels
 no contrastive vowel length
 phonemic stress present

The system that resulted from intrafamilial contact between Arabized Berber troops and Roman women was (5), which is still observable in northern Morocco (especially Tetuan and Chaouen), and in some Arab villages in the southern Rif.

- (5) Archaic MA
 {u a i} full vowels
 ə in unstressed posttonic syllables
 no contrastive vowel length
 phonemic stress present

From the perspective of the Berberized Arabic in (3), what happened is the following. First, the inherited vowel-length opposition collapsed. Those short vowels that had escaped syncope were merged with full vowels {u a i}. The original Arabic *ā/ī opposition had previously been reduced to *ə. This now had to be re-split, the choice between *a and *i depending chiefly on adjacent consonants (pharyngealized and back consonants requiring *a, otherwise *i). Old diphthongs *{āy āw} were either monophthongized to *{i u}, or lengthened to *{ay aw}. The merger of long and short vowels had dramatic consequences for archaic MA morphology, especially verbal morphology.

Second, a partially phonemic stress system developed. Stress is final on unsuffixed stems, but suffixation reveals limited stress oppositions, as in *salī-t* 'I finished' versus *sāl-ət* 'she questioned'. A vowel in a posttonic closed syllable (third feminine singular *-ət* is possibly the only example) is realized as schwa, which does not occur elsewhere in these archaic dialects.

These vocalic and prosodic changes were typical not only of the earliest MA in the Roman triangle (Tangier, Salé, Volubilis), but also very likely of early varieties of Spanish Arabic. Indeed, Corriente's analysis (e.g., 1997) of the later more mature Andalusī Arabic also recognizes merger of vowel length and the existence of stress contrasts.

Thereafter

The later history of MA was shaped by two phenomena: the progressive integration of local Berbers and Arabs in Morocco, and the Hilalian immigration that began in Tunisia. These two developments intersected in interesting ways.

The Roman triangle was somewhat insulated from Berber influence at the time of the Arab occupation. There was some interaction between Romans in Volubilis with Berbers, reaching into the nearby Middle Atlas. There was probably very little such interaction in Salé, and not much more in Tangier, where the bulk of the Roman population lived. This insulation did not change much in the first century of Arab rule, as the occupiers remained in the old Roman towns and cities, except for occasional raids on Berber populations to the south.

The Hilalians entered the Maghreb in Tunisia in the eleventh century. Over the next two centuries, they did two things of relevance to Morocco. First, they occupied much of the northern and western Sahara, as ancestors to present-day Hassaniya speakers in southern Algeria, northern Mali, Mauritania, the Western Sahara, and the southern oases of traditional Morocco including Guelmim and Tata. This led directly to a limited amount of contact with Pre-Hilalian Arabic speakers, as when a tribe of Hassaniya speakers from the southern oases were transported to the outskirts of Rabat by a Sultan to protect his flank.

However, the biggest effect of the Hilalian expansion was dialect mixing, initially in Algeria, resulting in hybrid dialects with a mix of pre-Hilalian and Hilalian features. These varieties were simultaneously deeply affected, especially in phonology and prosody, by Berber languages. Immigration from Algeria and Tunisia, mainly through the Taza corridor between Rif and Middle Atlas and on to the agriculturally rich plains in the Taza-Fes-Meknes-Rabat region, brought these varieties to central Morocco. These Arabs also became the core population of Meknes as it grew, and later combined with Berbers in Marrakesh. Additional Berber-Arabic contact has continued in these locations down to the present.

The Moroccan koiné is such a hybrid dialect. The mixing is phonologically observed in the q/g split. The mixing is epitomized in the competition between *dya*l ~ *d-* (of local Latin origin) and *ntaʃ* (non-Moroccan Maghreb, including Hassaniya) as genitive linkers, and by their interaction: the specifically feminine *ntaʃ-ət* and plural *ntawəʃ* have served as analogical models for feminine *dya*l-(ə)*t* and plural *dya*wə*l* in hybrid (but not archaic) MA. The main features of koiné vocalism and prosody are summarized in (6).

- (6) Hybrid MA (koiné)
 {u a i} full vowels
 ə short vowel
 phonemic stress absent

In most MA koiné dialects, seemingly, a second short (pseudo-)vowel occurs as in *xūbz* “bread” and *qūl* “kill”. However, these are analyzable phonemically as cases of

ə adjacent or nearly so to a velar or uvular that is associated with a labialization feature, hence structural /x^wəbz/ and /q^wtəl/. The labialization generally remains under syncope, as in *x^wbz-a* “(a) loaf of bread” and *q^wtl-at-u* “she killed him”. Only in dialects where ʔ occurs with a wider range of flanking consonants can this vowel be considered phonemic.

As is well known, hybrid dialects, especially in Morocco, also extend syncope farther than in archaic dialects under continuing Berber influence: e.g., archaic (northern) *daxl-u* “they entered” corresponds to koiné *dx^wl-u*.

Hereafter

As an MA specialist with limited experience in Berber linguistics, other than Tamashek (Tuareg of Mali), it would be interesting to hear a discussion by competent Berberists, especially about the likely vocalism of seventh–eighth century Berber varieties from Tunisia across to Morocco.

References

- Corriente F (1997) A dictionary of Andalusī Arabic. Brill, Leiden and New York
- Harrell R (1966) A dictionary of Moroccan Arabic, Arabic-English. Georgetown University Press, Washington
- Heath J (2015) D-possessives and the origins of Moroccan Arabic. *Diachronica* 32(1):1–33. <https://doi.org/10.1075/dia.32.1.01hea>
- Heath J (2018) Vowel-length merger and its consequences in early Moroccan Arabic. *Zeitschrift für arabische Linguistik* 67:12–43. <https://doi.org/10.13173/zeitarabling.67.0012>
- Marçais W (1911) *Textes arabes de Tanger: Transcription, traduction annotée, glossaire*. Leroux, Paris

Conjugated Adjectives and Participles in Senhaja Berber (Northwestern Morocco)

25

Evgeniya Gutova

Contents

Introduction	524
Adjectives and Participles in Senhaja	525
Adjectives as a Distinct Morphological Class	527
Conjugation of Adjectives and Participles in Senhaja	530
Introduction	530
Conjugated Berber Adjectives	531
Arabic Adjectives and Participles	533
Reference to the Future and Past	534
Negation of (Conjugated) Adjectives and Participles	536
Negation of (Conjugated) Adjectives and Participles in the Present	536
Negation of (Conjugated) Adjectives and Participles in the Future and Past	537
Relative Forms of Berber Adjectives, Arabic Adjectives, and Participles	538
Introduction	538
Berber Adjectives in Relative Constructions	539
Arabic Adjectives and Participles in Relative Constructions	541
Summary and Conclusion	543
References	544

Abstract

In Senhaja Berber (northwestern Morocco), unlike in many other Berber varieties, there is a distinct morphological class of adjectives, with two subclasses: native Berber, and borrowed Arabic adjectives. Adjectives have some common features with borrowed Arabic participles. Adjectives and participles behave differently across Senhaja varieties. In the varieties of Ketama and Hmed, adjectives are never conjugated, while in the Seddat variety, lexemes cognate to native Berber adjectives are obligatorily conjugated with the regular affixes, and are distinguished from verbs only by the lack of aspectual distinctions. In Taghzut, Bunsar, and Zerqet varieties, both conjugated and nonconjugated adjectives are found. Conjugated participles are

E. Gutova (✉)
University of Navarra (ICS), Pamplona, Spain

found in Taghzut and Bunsar. Conjugation of adjectives may be special, different from the verbs (Taghzut, Bunsar, and parts of Zerqet), or regular (parts of Zerqet). Conjugated adjectives can be regarded as quasi-verbs – an intermediate category between adjectives and verbs. Adjectives differ from verbs in that they mark only gender and number (and not person) and in the lack of aspectual distinctions. Conjugated adjectives mark the person but still lack aspectual distinctions.

Keywords

Conjugation · Adjectives · Participles · Conjugated adjectives · Word classes/ parts of speech distinction · Senhaja Berber

Abbreviations

1,2,3	1, 2, 3 person
A	Aorist
ADJ	Adjective
B	Bunsar
DO	Direct Object
F	Feminine
FUT	Future
H	Hmed
IMP	Imperative
IPF	Imperfect(ive)
IRR	Irrealis
K	Ketama
M	Masculine
NEG	Negation
P(L)	Plural
PFV	Perfective
PNG	Person, number, and gender
PRED	Predicative particle <i>ǝ</i>
PROX	proximal clitic
PST	Past marker
PTC	Participle
S(G)	Singular
T	Taghzut
VC	Ventive clitic
Z	Zerqet

Introduction

This chapter discusses the morphological classes of adjectives and participles, which can be conjugated in some varieties of Senhaja Berber. Senhaja is spoken by the Senhaja de Sraïr tribe confederation in the western part of the Rif mountains. This tribe confederation consists of ten tribes, which correspond to ten

subvarieties of Senhaja. Among the varieties discussed in this chapter, Ketama, Taghzut, and Seddat are Western, Senhaja, Hmed, and Bunsar are Central Senhaja, and Zerqet is Eastern Senhaja (see the map at the end of the chapter). On the basis of Moroccan census data from 2014 (HCP), the number of Senhaja speakers can be roughly estimated at ca. 85,000. Most Senhaja speakers are bilingual, speaking both Berber and dialectal (local) Moroccan Arabic (“Darija”), and code-switching is common.

The main publications on Senhaja are the following: a descriptive grammar by Renisio (1932), partly covering Senhaja; a vocabulary by Ibáñez (1959); a linguistic atlas by Lafkioui (2007); and an article by Kossmann (2017), discussing the place of Senhaja within Berber. While Senhaja is geographically close to other varieties of the Rif (Tarifit), it shows some unique linguistic features (Kossmann 2017; Lafkioui 2007; Gutova 2021, 2022), distinguishing it from the varieties spoken to the East (Tarifit) and to the South (Central Atlas Tamazight). It shares some features with Ghomara Berber spoken to the West (El Hannouche 2008; Mourigh 2016).

The present chapter consists of the following parts: Section “[Adjectives and Participles in Senhaja](#)” provides an overview of adjectives and participles in Senhaja, discusses their defining features and morphological markers. Section “[Adjectives as a Distinct Morphological Class](#)” devotes more attention to adjectives as a distinct morphological class and compares expression of qualities in other Berber languages. Section “[Conjugation of Adjectives and Participles in Senhaja](#)” discusses the conjugation of adjectives and participles in Senhaja. For the reference to the future and past, see section “[Reference to the Future and Past](#).” Section “[Negation of \(Conjugated\) Adjectives and Participles](#)” discusses the negation strategies. Adjectives and participles in contexts of relativization are discussed in section “[Relative Forms of Berber Adjectives, Arabic Adjectives, and Participles](#).” The last section provides a summary and draws some conclusions.

In this chapter, IPA is used rather than the transcription system that is common in literature on Berber. Consider the following IPA symbols and their equivalents in Berberological literature, listed between <...>: \hbar <ḥ>, ζ <ε>, \mathfrak{f} <š> or <c>, \mathfrak{z} <ž>, $\mathfrak{t}\mathfrak{f}$ <č>, $\mathfrak{d}\mathfrak{z}$ <ğ>, j <y>. Also, in Berberological transcription, long consonants are doubled, e.g., b : <bb>, pharyngealized consonants are written with a dot underneath, e.g., $t^\text{ʕ}$ <ṭ>, and spirantized consonants are written with an underscore, e.g., θ <ṯ>, δ <ḍ>, $\delta^\text{ʕ}$ <ḍ̣>, β <ḅ>, and ς <ḳ>. The symbol ^ in this chapter indicates elision and assimilation

Adjectives and Participles in Senhaja

In Senhaja, adjectives constitute a word class of their own, with defined features. There are native Berber and borrowed Arabic adjectives, with different morphological markers, but the same function. The class of Berber adjectives is limited (there are only six lexemes), while Arabic adjectives are numerous (103 examples in our database). The Berber adjectives are as follows: $m\acute{a}z:i/m\acute{a}z^\text{ʕ}:i$ “small” (most Senhaja except Hmed); $m\acute{a}\mathfrak{f}:i\varsigma$ “small” (Hmed); $m\acute{a}q:(u)r$ “big” (pan-Senhaja); and $m\acute{a}l:uj$

“white” (Hmed); *məs:us* “insipid” (Hmed); and *rzaj* “bitter” (Hmed). There are dialectal differences in the frequency of Arabic borrowed adjectives. While Ketama often employs Arabic adjectives without any change in their morphology, in Zerqet, these are often incorporated into the (regular) verbal class. An adjectival class is most likely an innovation in Senhaja: Berber adjectives have developed from the old stative verbs, while Arabic adjectives are a result of borrowing. A similar situation is found in Ghomara (El Hannouche 2008; Mourigh 2016). Adjectives usually modify nouns but can also function as predicates. They obligatorily agree with the head noun in gender and number. The adjective can be nominalized. In certain contexts, Arabic adjectives accept the Arabic article *l-*.

Participles have also been borrowed from Arabic and form a distinct word class in Senhaja. Participles borrowed from Arabic are distinct from what is traditionally known as “the Berber participle,” which is a special form of the verb used in contexts of relativization (here referred to as “the relative form,” to avoid confusion with participles borrowed from Arabic); see section “[Relative Forms of Berber Adjectives, Arabic Adjectives, and Participles](#)”. Arabic participles are also found in Ghomara (Mourigh 2016). In Arabic, participle formation is productive. In Senhaja, Arabic participles are borrowed as such, rather than formed within Berber. Exceptionally, some participles are formed following the Arabic pattern based on the native Berber verbs. This phenomenon is most frequent in Hmed, while a few examples are found in Ketama and Taghzut. There is no separate class of Senhaja native participles that differ from Arabic participles in marking. When nominalized, participles can also function as nouns. Senhaja employs a large number of Arabic participles (220 in our database), the majority of which are passive participles (144 in our database, as opposed to 76 active participles). Almost all participles derived from native Berber verbs are passive. Not all participles are found across all Senhaja varieties. However, it is safe to say that all Senhaja varieties have borrowed a vast number of Arabic participles and use them in everyday speech.

Participles retain Arabic morphology in Senhaja. The difference is that in Arabic, participles are derived from the corresponding verb root by applying a certain scheme, while in Senhaja, many participles do not share the root with the corresponding verb. Instead, they often have a suppletive relation to the native Berber verb. A given participle may be shared across Senhaja, while corresponding to different verbs in different varieties.

Adjectives and participles share a number of features in Senhaja:

- They can modify the head noun (see examples below).
- They distinguish three forms: masculine singular, feminine singular, and plural, with the possibility to (optionally) mark feminine plural in Zerqet. (The morphemes of gender and number are not the same as with nouns. Also, nouns obligatorily have a distinct FP form, while adjectives have a common plural form in most Senhaja varieties.)
- Arabic adjectives with external plurals and participles share the same suffixes to mark the gender and number.

Table 24.1 Adjectival and participle marking in Senhaja

	Berber adjectives			Arabic adjectives and participles		
	sfx	Berber ADJ “small”		sfx	Ar. ADJ “bad”	PTC “washed”
		Ketama	Zerqet		Ketama/Zerqet	Ketama/Zerqet
MS	Ø	məz:i	məz ^ɕ :i	Ø	ʕəj:an	məysul
FS	-(ə)θ	məz:i-θ	məz ^ɕ :ij-əθ	-a	ʕəj:an-a	məysul-a
MP	-(ə)n	məz:i-n	məz ^ɕ :ij-ən	-in	ʕəj:an-in	məysul-in
FP	-(ə)n(t)	=MP	məz ^ɕ :ij-ən(t)	-aθ (Z)	ʕəj:an-aθ (Z)	məysul-aθ (Z)

The following two examples show the use of an adjective (*məq:ur^ɕ* “big”) and a participle (*məsbuy* “painted”) as noun modifiers:

- (1) *k:r-ən* *iməħzar*, *d:a-n* *ħət:a[^]* (*a*)ni *rəkm-ən*
 get.up:P-3PL orphans go:P-3PL until where reach:P-3PL
ar *jan* *wərθi* *məq:ur^ɕ* (K)
 to one field:MS big:MS
 “The orphans got up, and walked until they reached a big field.”
- (2) *zug:ər* *ax:am =aθ* *məsbuy* (K)
 look.IMP.SG house =PROX:SG painted:MS
 “Look at this painted house.”

Table 24.1 shows the morphological markers used with Berber adjectives (such as *məz:i~məz^ɕ:i* “small”), Arabic adjectives (such as *ʕəj:an* “tired”), and borrowed Arabic participles (such as *məysul* “washed”) in Senhaja. The word for “small” shows variation with regard to pharyngealization (*z:* in Ketama/Taghzut/Bunsar, *z^ɕ:* in Seddat/Zerqet) and the final *-j* (present in Bunsar and Zerqet when followed by a suffix).

There are also some differences between adjectives and participles:

- There are differences in the possibility to conjugate adjectives and participles (also depending on the variety).
- There are differences in the use and marking of adjectives and participles in relative constructions: Berber adjectives have a special relative form used in contexts of relativization, while Arabic adjectives and participles do not.
- Adjectives (but not participles) allow for the formation of diminutives in some Senhaja varieties.

Adjectives as a Distinct Morphological Class

A separate morphological class of adjectives is not found in all Berber varieties. There is a debate in the literature on Berber about adjectives as belonging to a specific lexical category, or rather being a subcategory of nouns or verbs

(see Chaker 1985: 1–2; Galand 2002: 199; Taine-Cheikh 2009). There is also a debate about whether it is a subcategory at all. See Galand 1990 on the diachronic development from nouns to stative verbs (*verbe d'état*) in Berber. Taine-Cheikh (2003) discusses conjugated adjectives in Zenaga, while Taine-Cheikh (2014) discusses the verb-noun distinction in Berber and its fluctuations.

Expression of noun attributes is carried out by various means in Berber languages, such as the use of verbs (regular or stative) or attribute nouns that can be used to modify the head noun. For example, native Berber color terms are attribute nouns in Senhaja. They are used alongside borrowed Arabic adjectives in most Senhaja varieties, e.g., (Hmed) *aβərçan* (a native Berber attribute noun) “a black one” is used alongside *kħəl* (a borrowed Arabic adjective) “black.” The predicative (copular) particle *δ* is optionally used with nominal attributes in Hmed (and obligatorily in Zerqet) but is not used with (Berber or Arabic) adjectives when used as predicates. In Ketama, the predicative *δ* is not used in any context. Contrast the following examples (Hmed), where a native Berber attribute noun is used in (a), with an optional copula, and a borrowed Arabic adjective in (b), where the use of a copula is impossible:

- (3) (a) *axam inu (δ) aβərçan* (H)
 house:MS my PRED black.one:MS
 (b) *axam inu kħəl* (H)
 house:MS my black:MS
 “My house is black.”

In (a), the noun *aβərçan* “a black one” modifies the noun *axam* “house.” Words such as *aβərçan* “a black one” are regarded as a subcategory of nouns, because they have morphological characteristics of usual nouns, and because not all nouns can be used attributively.

In many Berber languages, there is a distinct class of stative verbs. Stative verbs have special morphological markers (suffixes only) for expressing person, number, and gender (henceforth: PNG), distinguishing them from the regular verbs that use both prefixes and suffixes to mark PNG (cf. Kossmann 2009). In most Senhaja varieties, ancient stative verbs developed into adjectives that function syntactically in the same way as adjectives borrowed from Arabic, although they bear different morphological markers (cf. Table 24.1 above). There is thus a separate class of adjectives, with two subcategories: adjectives of Arabic origin, which are numerous, and adjectives of Berber origin, which are very few in number.

Lafkioui (2009: 111) analyzes Berber adjectives in Senhaja as verbs. According to Lafkioui, in the Bshir Senhaja variety, such lexemes have either the special conjugation or the regular conjugation, depending on the syntactic context: The special conjugation is found in combination with another verb (e.g., “to become,” as in 4a, and the regular conjugation is found when the form is used on its own, not as a complement of a verb as in 4b). The following two examples in Bshir variety are taken from Lafkioui 2009: 111; the gloss is added by the present author:

- (4) (a) θ -*ayul*- $\text{\textcircled{a}}\delta$ *məq:ur*
 2S-become:P-2S big:MS
 “You became big/old (MS).”
 (b) θ -*məq:ur*- $\text{\textcircled{a}}\delta$
 2S-big-2S
 “You (SG) are big/old.”

We have not observed this situation in the surveyed Senhaja varieties. In Taghzut and Bunsar, where variation in the marking of adjectives is found, both conjugated and nonconjugated forms are found in the same syntactic contexts, e.g., (Taghzut) (The word *məq:ur^s* shows variation with regard to the status of *u*. It is unstable in Taghzut and Seddat: *məq:ur^s* > *məq:r^s* when followed by a suffix.)

- (5) (a) θ -*uyul*- $\text{\textcircled{a}}\delta$ *məq:ur^s* (Taghzut)
 2-become:P-2S big:MS
 “You became big/old (MS).”
 (b) θ -*uyul*- $\text{\textcircled{a}}\delta$ *məq:r^s*- $\text{\textcircled{a}}\delta$ (Taghzut)
 2-become:P-2S big-2S
 “You (SG) became big/old.”
 (6) (a) $\text{\textcircled{c}}\text{\textcircled{a}}\text{\textcircled{d}}\text{\textcircled{z}}:i$ *məq:ur^s* (Taghzut)
 you:MS big:MS
 (b) ($\text{\textcircled{c}}\text{\textcircled{a}}\text{\textcircled{d}}\text{\textcircled{z}}:i$) *məq:r^s*- $\text{\textcircled{a}}\delta$ (Taghzut)
 (you:MS) big-2S
 “You (MS) are big/old.”

Such lexemes can be conjugated (with the regular or special PNG markers) only in parts of Senhaja, while in other parts (Ketama, Hmed), they are never conjugated. In Zerqet, conjugation depends on the person of the verb and the dialect (cf. below). When not conjugated, there are only three forms (cf. Table 24.1): MS (no affixes), FS (suffix *-θ*), and one common plural form (suffix *-n*), with an optional distinct FP form in Zerqet, which parallels the distinction in borrowed Arabic adjectives. We analyze such lexemes as adjectives, while noting that these adjectives can be conjugated in some varieties, whether with special or regular markers.

Table 24.2 summarizes the variation in Senhaja in the marking of Berber adjectives, taking as example the phrase “You are small/young.” In (a), a nonconjugated

Table 24.2 Marking of Berber adjectives in Senhaja: “you are young”

	Ketama	Taghzut	Seddat	Hmed	Bunsar	Zerqet
a) No person marking	$\text{\textcircled{c}}\text{\textcircled{a}}\text{\textcircled{d}}\text{\textcircled{z}}:i$ $məz:i$	$\text{\textcircled{c}}\text{\textcircled{a}}\text{\textcircled{d}}\text{\textcircled{z}}:i$ $məz:i$	—	$\text{\textcircled{c}}\text{\textcircled{a}}\text{\textcircled{d}}\text{\textcircled{z}}:i$ $mətf:i\text{\textcircled{c}}$	$\text{\textcircled{c}}\text{\textcircled{a}}\text{\textcircled{d}}\text{\textcircled{z}}:i$ $məz:i$	$\text{\textcircled{c}}\text{\textcircled{e}}\text{\textcircled{z}}\text{\textcircled{z}}\text{\textcircled{z}}\text{\textcircled{z}}:i$
b) Special PNG	—	$məz:i-\text{\textcircled{a}}\delta$	—	—	$məz:ij-\text{\textcircled{a}}\delta$	$məz^\text{\textcircled{c}}:ij-\text{\textcircled{a}}\delta$
c) Regular PNG	—	—	θ - $məz^\text{\textcircled{c}}:i-\text{\textcircled{a}}\delta$	—	—	θ - $məz^\text{\textcircled{c}}:ij-\text{\textcircled{a}}\delta$

adjective is used with the 2MS independent pronoun $\zeta\hat{a}d\zeta:i$ (nonconjugated adjectives are found in most Senhaja varieties, except for Seddat), while in (b) and (c), the conjugated forms are used, and the gender is unmarked. In Zerqet, the three forms are not found in the same variety but reflect dialectal preferences (cf. below for more details). The word for “small” is $m\acute{a}z:i$ or $m\acute{a}z^{\epsilon}:i(j)$ in most Senhaja varieties, while Hmed has $m\acute{a}f:i\zeta$.

As follows from the table, in Ketama and Hmed, adjectives are never conjugated, while in Seddat, cognate lexemes are always obligatorily conjugated with the regular PNG affixes. For Seddat, such words are best analyzed as defective verbs that are distinguished from regular verbs only by the lack of aspectual distinctions. Finally, in some varieties (Taghzut, Bunsar, and Zerqet), both conjugated and nonconjugated adjectives are found.

Conjugation may be carried out by suffixes only (Taghzut, Bunsar, and parts of Zerqet) or can be regular (parts of Zerqet). Conjugated adjectives (Taghzut, Bunsar, and Zerqet) can be regarded as an intermediate category between adjectives and verbs. There is thus a continuum between nonverbs and verbs in Senhaja: adjectives differ from verbs in that they mark only gender and number (and not person) and in that they have no aspectual distinctions. Conjugated adjectives mark the person in addition to number and gender, which makes them similar to verbs, although they lack aspectual distinctions. In Seddat, there are no Berber adjectives, but only defective verbs. In varieties that have suffixal conjugation (Taghzut, Bunsar, and Zerqet), not all verbs describing attributes have a special conjugation, and some attributive verbs have a regular marking.

Conjugation of Adjectives and Participles in Senhaja

Introduction

As discussed above, some Senhaja varieties (Taghzut, Bunsar, and Zerqet) allow for Berber adjectives to be conjugated. As mentioned above, in Seddat, cognate lexemes are obligatorily conjugated and are regarded as defective verbs, while in Ketama and Hmed, adjectives are not conjugated. Ghomara Berber also does not have conjugated adjectives (Mourigh 2016). Furthermore, in Taghzut and Bunsar, but not in Zerqet, Arabic adjectives and participles can also be conjugated. The adjectival conjugation found in Taghzut, Bunsar, and parts of Zerqet is carried out by suffixes only, and as such, it recalls the stative conjugation found in many Berber varieties (Kossmann 2009). However, we do not analyze conjugated adjectives as stative verbs, because they lack the aspectual distinctions. Following Taine-Cheikh (2003, 2014), we analyze such forms as conjugated adjectives. In Zenaga, adjectives are also conjugated by suffixes only. In Zenaga, for conjugated adjectives, there is often also a Perfective of the same root that has a regular conjugation. In Senhaja, the Berber adjective $m\acute{a}q:ur^{\epsilon}$ “big” (which can be conjugated in parts of Senhaja) exists alongside a cognate dynamic verb $myur^{\epsilon}$ “to be(come) big/old, to grow (up)” (y corresponds to q : when it is long). The Berber adjective $m\acute{a}z:i/m\acute{a}z^{\epsilon}:i$ “small” does not have a corresponding

dynamic verb in most Senhaja varieties. However, in Zerqet, the adjective *məzːi* “small” has a corresponding verb *mzːi* “to be(come) small(er).” Note that here as well, a short consonant in the verb corresponds to a geminated consonant in the corresponding adjective, as is the case with *mɣur* “be(come) big/old, grow up” > *məqːur* “big/old.” Consider the following examples:

- (7) *f˘ad* *i-mzːi* (Z)
 FUT^IRR 3MS-be.small.A
 “He/It (M) will be(come) small(er).”
- (8) *i-mzːi* (Z)
 3MS-be.small.PFV
 “He/It (M) became small(er).”
- (9) *kul* *nːhar* *i-təmzːijaj* (Z)
 every day 3MS-be.small.IPF
 “Every day he gets smaller.”

The verbs *mɣur* (pan-Senhaja) “to be(come) big/old” and *mzːi* (Zerqet) “to be (come) small” can be used in all aspects (Aorist, Perfective, and Imperfective), and thus, the adjectives *məqːur* “big/old” and *məzːi/məzːi* “small” do not replace the Perfective forms of these verbs. In this regard, Senhaja is similar to Zenaga. Contrast the following examples (pan-Senhaja), with an adjective describing a state in (10a), and a dynamic verb describing a result in (10b):

- (10) (a) *Muħəmːəð* *məqːur*
 Mohammed big:MS
 “Mohammed is big.”
- (b) *Muħəmːəð* *i-mɣur*
 Mohammed 3MS-grow.PFV
 “Mohammed grew (up).”

Conjugated adjectives are akin to defective verbs and are distinguished from them in Taghzut and Bunsar by the special marking. In Seddat, the conjugation of cognate lexemes is regular, making them defective verbs. The line between the conjugated adjectives and defective verbs is thus drawn based on morphological criteria. In what follows, we first discuss conjugated Berber adjectives in Taghzut and Bunsar with the cognate lexemes in Seddat and Zerqet (Section “[Conjugated Berber Adjectives](#)”). Conjugated Arabic adjectives and participles (Taghzut and Bunsar) are discussed in section “[Arabic Adjectives and Participles](#).” In this section, only forms in affirmative contexts and with reference to the present situation are discussed.

Conjugated Berber Adjectives

In parts of Senhaja where Berber adjectives (or cognate lexemes) can be conjugated, there are two sets of PNG markers: the *regular set* (the same as found with verbs,

constituted by prefixes and suffixes), and the *special set* (constituted by suffixes only). In Seddat, cognate lexemes (regarded here as defective verbs) are obligatorily conjugated with the regular PNG affixes. In Zerqet, the situation is special (see below).

Table 24.3 shows the paradigm of conjugated adjectives and Seddat defective verbs. Empty cells in the table indicate the absence of conjugation markers. In some forms (3MS and 3FS in the special set), there is no formal difference in the marking of nonconjugated and conjugated adjectives. In 2P, Bunsar has *-n*, Zerqet *-m*, while Taghzut shows variation (*-n~-m*). Zerqet (the Ikherruden dialect) optionally distinguishes separate 2FP and 3FP forms (absent in Taghzut, Bunsar, and Seddat).

In Zerqet, different markings are found in different dialects (Bunjel and Ikherruden). In Bunjel, only suffixes are used (*special set*). In Ikherruden, the paradigm is similar to the regular one, except for 3FS, which can combine the prefix *θ-* and the suffix *-θ* (the suffix can be omitted). Unlike Taghzut and Bunsar, Zerqet does not allow the use of nonconjugated or conjugated adjectives in all persons. In Zerqet, depending on the variety, nonconjugated (Bunjel) or conjugated (Ikherruden) forms are found for the 1S. In 2P, the person is obligatorily marked in Zerqet (Bunjel *-m*, Ikherruden *θ-...-m*), while a common plural form found in other Senhaja varieties (*-n*) is not used. Table 24.4 shows paradigms for the conjugated adjective *məq:ər^ɕ~məq:ur^ɕ* “big, old” and Seddat defective verb “to be big/old.” Table 24.5 illustrates the conjugation of the adjective *məz:i~məz^ɕ:i* “small, young” and the Seddat defective verb *məz^ɕ:i* “to be small/young.”

In the following example, the conjugated adjective is used predicatively:

- (11) *nək məq:ur^ɕ-aɣ* (Z-Ikherruden)
 1S big-1S
 “I am big/old.”

Table 24.3 Conjugation of adjectives and defective verbs in Senhaja

Taghzut/Bunsar/Zerqet-Bunjel (<i>special set</i>)			Seddat (<i>regular set</i>)			Zerqet-Ikherruden (<i>regular set</i>)				
1S	X	-(a)ɣ ^a	1S		X	-(a)ɣ	1S		X	-(a)ɣ
2S	X	-(ə)ð	2S	θ-	X	-(ə)ð	2S	θ-	X	-(ə)ð
3MS	X		3MS	i-	X		3MS	i-	X	
3FS	X	-(ə)θ	3FS	θ-	X		3FS	θ-	X	-(ə)θ
1P	X	-(ə)n	1P	n-	X		1P	n-	X	
2P	X	-(ə)n (B/T) -(ə)m (T/Z)	2P	θ-	X	-(ə)m	2P	θ-	X	-(ə)m
2FP			2FP				2FP	θ-	X	-(ə)mt
3P	X	-(ə)n	3P		X	-(ə)n	3P		X	-(ə)n
3FP			3FP				3FP		X	-(ə)nt

^aIn the Bunjel dialect of Zerqet, only nonconjugated adjectives are found with 1S, with the gender marking (masculine X, feminine X-θ)

Table 24.4 Conjugation of *məq:(u)r^c* “(to be) big/old”

Taghzut/Bunsar/Zerqet-Bunjel (<i>special set</i>)			Seddat (<i>regular set</i>)			Zerqet-Ikherruden (<i>regular set</i>)		
1S	məq:(u)r ^c	-aɣ ^a		məq:r ^c	-aɣ		məq:ur ^c	-aɣ
2S	məq:(u)r ^c	-əð	θ-	məq:r ^c	-əð	θ-	məq:ur ^c	-əð
3MS	məq:(u)r ^c		i-	məq:ur ^c		i-	məq:ur ^c	
3FS	məq:(u)r ^c	-əθ	θ-	məq:ur ^c		θ-	məq:ur ^c	(-əθ)
1P	məq:(u)r ^c	-ən	n-	məq:ur ^c		n-	məq:ur ^c	
2P	məq:(u)r ^c	-ən (B/T) -əm (T/Z)	θ-	məq:r ^c	-əm	θ-	məq:ur ^c	-əm
2FP						θ-	məq:ur ^c	-əmt
3P	məq:(u)r ^c	-ən		məq:r ^c	-ən		məq:ur ^c	-ən
3FP							məq:ur ^c	-ənt

^aIn Bunjel, only nonconjugated adjectives are found with 1S (MS *məq:ur^c*, FS *məq:ur^c-əθ*)

Table 24.5 Conjugation of *məz:i(j)~məz^c:i(j)* “(to be) small”

Taghzut/Bunsar/Zerqet-Bunjel (<i>special set</i>)			Seddat (<i>regular set</i>)			Zerqet-Ikherruden (<i>regular set</i>)		
1S	məz:i(j)	-aɣ ^a		məz ^c :i	-ɣ		məz ^c :ij	-aɣ
2S	məz:i(j)	-əð	θ-	məz ^c :z ^c :i	-ð	θ-	məz ^c :ij	-əð
3MS	məz:i		i-	məz ^c :i		i-	məz ^c :i	
3FS	məz:i(j)	-əθ	θ-	məz ^c :i		θ-	məz ^c :i(j)	(-əθ)
1P	məz:i(j)	-ən	n-	məz ^c :i		n-	məz ^c :i	
2P	məz:i(j)	-(ə)n (B/T) -(ə)m (T/Z)	θ-	məz ^c :i	-m	θ-	məz ^c :ij	-əm
2FP						θ-	məz ^c :ij	-əmt
3P	məz:i(j)	-ən		məz ^c :i	-n		məz ^c :iy	-ən
3FP							məz ^c :iy	-ənt

^aIn Bunjel, only nonconjugated adjectives are found with 1S (MS *məz^c:i*, FS *məz^c:ij-əθ*)

In the following example, the conjugated adjective is used as a modifier:

- (12) *i-w:i* =d *aʕək^w:az* *i-məz^c:i* (Z-Ikherruden)
 3MS-bring.PFV =VC stick 3MS-small
 “He brought a small stick.”

Arabic Adjectives and Participles

In Taghzut and Bunsar, borrowed Arabic adjectives and participles can also be conjugated. The PNG marking is similar, but not identical to the one found with Berber adjectives as discussed above. The marking is different in that there is an extra vowel *i* before the PNG suffixes. In the 2P, Taghzut shows variation (*-im/-in*),

Table 24.6 Conjugation of Arabic adjectives and participles in Taghzut and Bunsar

	Scheme		ADJ <i>məzjan</i> “good”		PTC <i>sakən</i> “dwelling”		PTC <i>mafi</i> “going”	
1S	X	-iy	məzjan	-iy	sakn	-iy	maʃj	-iy
2S	X	-ið	məzjan	-ið	sakn	-ið	maʃj	-ið
3MS	X		məzjan		sakən		maʃi	
3FS	X	-a	məzjan	-a	sakn	-a	maʃj	-a
1P	X	-in	məzjan	-in	sakn	-in	maʃj	-in
2P	X	-in (T/B) -im (T)	məzjan	-in, -im	sakn	-in, -im	maʃj	-in, -im
3P	X	-in	məzjan	-in	sakn	-in	maʃj	-in

while in Bunsar, only *-in* is found. Table 24.6 shows the scheme of the conjugation and the examples: the Arabic adjective *məzjan* “good,” and the participles *sakən* “dwelling” and *mafi* “going.” (In Zerqet, the participle *mafi* “going” has a geminated *f*.) When the stem of an Arabic adjective or participle ends on *-i*, as in *mafi* “going,” this *i* becomes *j* before suffixes: *maʃj*-*iy* (T/B) “I am going.”

Reference to the Future and Past

When adjectives and participles appear on their own (whether conjugated or not), they refer to the situation in the present, or no specific time reference is made. In Taghzut, the use of the verb “to be” (in Perfective) is optional in this case. The future is expressed by means of the marker $(ma)ʃ^a(\delta)$ + the verb “to be” in the Aorist. (The marker $(ma)ʃ^a\delta$ historically consists of the future marker $(ma)ʃ$ (ultimately from the borrowed Arabic participle *mafi* “going”) followed by the Berber irrealis marker $a\delta$. The Aorist stem of the verb “to be” is realized as *iʒi* in Taghzut and *ij:i* in Bunsar (Tamadit dialect) – both issuing from *ili*, as still found in Zerqet and in the Luta dialect of Bunsar. The Perfective form is realized as *l:i/l:a* in Taghzut, and as $\bar{d}\bar{z}:i/\bar{d}\bar{z}:a$ in Bunsar and Zerqet (with the vowel alternation).) The reference to the past is formed by the past marker (obligatory in Taghzut, optional in Bunsar/Zerqet) followed by the verb “to be” in the Perfective; the verb “to be” is optional in Taghzut. The past marker is *indi* in Taghzut, and *iza* in Bunsar and Zerqet (also *za* in Bunsar). (The past marker *indi* is found within Senhaja only in Taghzut.) The marker *iza* is a (3MS) petrified form of the verb “to be.” Similar strategies to refer to the future and past are used across Senhaja. (The same strategies are used to refer to the future and past with nonconjugated and conjugated forms of adjectives and participles. The difference is that conjugated forms mark the person, while non-conjugated forms allow only for gender and number marking.) Schematically:

1. Future: FUT $(ma)ʃ^a$ + “to be: Aorist” + ADJ/PTC
2. Present: (in Taghzut, optionally: the verb “to be: Perfective”) + ADJ/PTC
3. Past: PST marker + “to be: Perfective” (optional in Taghzut) + ADJ/PTC

Consider examples with the 1S conjugated form of the adjective *məq:(u)r^ε* “big/old” in Table 24.7.

In Bunsar and Zerqet, to refer to the future, there exists an alternative construction based on the verb *af* “to find,” e.g., (Zerqet)

- (13) *ʃ[^]a* *j=* *θ-af-əð* *məq:ur^ε*
 FUT[^]IRR 1S:DO= 2S-find.A-2S big:MS
 “I will be big/old(er),” lit. “You (SG) will find me big/old(er).”

The same construction is possible with a conjugated form of the adjective but is less common:

- (14) *ʃ[^]a* *j=* *θ-af-əð* *məq:ur^ε-ay*
 FUT[^]IRR 1S:DO= 2S-find.A-2S big-1S
 “I will be big/old(er),” lit. “You (SG) will find me big/old(er).”

Compare the construction with a cognate dynamic verb *myur* “to grow (up)”:

- (15) *ʃ[^]a* *j=* *θ-af-əð* *myur-ay*
 FUT[^]IRR 1S:DO= 2S-find.A-2S grow.PFV-1S
 “I will have grown (already),” lit. “You will find me having grown.”

The same strategies (the use of an auxiliary “to be” and the past marker) are used to refer to the situation in the future and in the past with conjugated forms of Arabic adjectives and participles in Taghzut and Bunsar. Consider examples with the 1S conjugated forms of the Arabic adjective *məzjan* “good” and participle *sakən* “dwelling” (Tables 24.8 and 24.9).

Table 24.7 Reference to the future, present, and past with a Berber adjective

Meaning	Taghzut	Bunsar/Zerqet	Gloss
1) I will be big	<i>ʃ[^]að iʒi-γ məq:r^ε-ay</i>	<i>ʃ[^]að ili-γ məq:ur^ε-ay^a</i>	FUT [^] IRR be.A-1S big-1S
2) I am big	<i>(l:i-γ) məq:r^ε-ay</i>	<i>məq:ur^ε-ay</i>	(be.PFV-1S) big-1S
3) I was big	<i>indi (l:i-γ) məq:r^ε-ay</i>	<i>(iʒa) dʒi-γ məq:ur^ε-ay</i>	PST be.PFV-1S big-1S

^aBunsar (dialectal): *ʃ[^]að ij:i-γ məq:ur^ε-ay*

Table 24.8 Reference to the future, present, and past with an Arabic adjective

Meaning	Taghzut	Bunsar (Tamadit)	Gloss
1) I will be good	<i>ʃ[^]að iʒi-γ məzjan-iy</i>	<i>ʃ[^]að ij:i-γ məzjan-iy</i>	FUT [^] IRR be.A-1S good-1S
2) I am good	<i>(l:i-γ) məzjan-iy</i>	<i>məzjan-iy</i>	(be.PFV-1S) good-1S
3) I was good	<i>indi (l:i-γ) məzjan-iy</i>	<i>(i)ʒa dʒi-γ məzjan-iy</i>	PST be.PFV-1S good-1S

Table 24.9 Reference to the future, present, and past with an Arabic participle

Meaning	Taghzut	Bunsar (Tamadit)	Gloss
1) I will be dwelling	ʃʌð iʒi-γ sakn-iγ	ʃʌð ij:i-γ sakn-iγ	FUT ^{IRR} be.A-1S dwelling-1S
2) I am dwelling	(l:i-γ) sakn-iγ	sakn-iγ	(be.PFV-1S) dwelling-1S
3) I was dwelling	indi (l:i-γ) sakn-iγ	(i)ʒa dʒi-γ sakn-iγ	PST be.PFV-1S dwelling-1S

Negation of (Conjugated) Adjectives and Participles

Negation of (Conjugated) Adjectives and Participles in the Present

The negation of adjectives and participles with reference to the present is achieved in one of the following ways:

1. By the negation of “to be”: u + “to be: Perfective” + $f/i/aj$ + ADJ/PTC. The postverbal negator across Senhaja can be f , f/i , or faj , which are usually in free variation.
2. By a single negator *mafi* (borrowed from Arabic): *mafi* + ADJ/PTC. In most Senhaja varieties, the negator *mafi* is homonymous with the participle *mafi* “going”. In Zerqet, the participle has a geminated f : *maf:i*.
3. By the bipartite negation: u + ADJ/PTC + $f(aj)$. Taghzut: $u \sim ma \sim u$ *ma* + ADJ/PTC + $f(aj)$. In Taghzut, the preverbal negator can be u (native Berber), *ma* (borrowed from Arabic), or the combination of the two: u *ma*. There is a preference to use the originally Arabic negator *ma* with the Arabic loans. There is also a detectable shift to the negator *ma* with the younger generation.

With the first and the second negation strategies, in Taghzut and Bunsar, both conjugated and nonconjugated adjectives can be negated in this way. In Zerqet, usually, nonconjugated adjectives are negated in these ways. The negation of conjugated adjectives with the help of the auxiliary verb “to be” or a single negator *mafi* is possible given the right context but is less common in Zerqet. The third strategy, by means of the bipartite (discontinuous) negation $u...f(aj)$, is specific to the conjugated adjectives in Bunsar and Zerqet, while the negation of nonconjugated adjectives by this means is considered ungrammatical. In Taghzut, there is variation: Some speakers accept the use of the bipartite negation with nonconjugated adjectives, while others do not. All our consultants come from the same village in Taghzut, Lqela, so the difference is not dialectal. It is also not caused by the age difference. Table 24.10 provides examples of the negated 1S conjugated form of the adjective *maq:(u)r* “big/old” (“I am not big/old”). Table 24.11 provides examples of the negated 1S conjugated forms of the Arabic adjective *məzjan* “good” and participle *sakən* “dwelling” (Taghzut/Bunsar).

In Taghzut, nonconjugated adjectives can be negated in the same way, including the third strategy with the bipartite negation. Compare the examples with a conjugated adjective in (16a), and a nonconjugated adjective in (16b).

Table 24.10 Negation of a Berber adjective: “I am not big/old”

	Taghzut	Bunsar	Zerqet (Ikherruden)	Gloss
1)	u l:i-γ f məq:r ^ε -aγ	u ḍɜ:i-γ f məq:ur ^ε -aγ	(u ḍɜ:i-γ f məq:ur ^ε -aγ) ^a	NEG be.PFV-1S NEG big-1S
2)	mafi məq:r ^ε -aγ	mafi məq:ur ^ε -aγ	(mafi məq:ur ^ε -aγ) ^b	NEG big-1S
3)	u (ma) məq:r ^ε -aγ f	u məq:ur ^ε -aγ f	u məq:ur ^ε -aγ f	NEG big-1S NEG

^aMore common: (*nək*) u ḍɜ:i-γ f məq:ur^ε (1S NEG be.PFV-1S NEG big:MS) “I am not big/old” (with a nonconjugated adjective)

^bMore common: *nək mafi məq:ur^ε* (1S NEG big:MS) “I am not big/old” (with a nonconjugated adjective)

Table 24.11 Negation of Arabic adjectives and participles

	ADJ <i>məzjan</i> “good”: “I am not good”		PTC <i>sakən</i> “dwelling”: “I am not dwelling”	
	Taghzut	Bunsar	Taghzut	Bunsar
1)	u l:i-γ f məzjan-iγ	u ḍɜ:i-γ f məzjan-iγ	u l:i-γ f sakn-iγ	u ḍɜ:i-γ f sakn-iγ
2)	mafi məzjan-iγ	mafi məzjan-iγ	mafi sakn-iγ	mafi sakn-iγ
3)	u (ma) məzjan-iγ f	u məzjan-iγ f	(u) ma sakn-iγ f	u sakn-iγ f

- (16) (a) *u ma məq:r^ε-aγ f(aj)*
 NEG NEG big-1S NEG
 (b) *nekki u ma məq:ur^ε f(aj)*
 I NEG NEG big:MS NEG
 “I am not big/old.”

Negation of (Conjugated) Adjectives and Participles in the Future and Past

The negation with the reference to the future is formed by negating the verb “to be” in the Aorist (with the future marker *f^{ad}*). The negation with the reference to the past is formed by negating the verb “to be” in the Perfective preceded by the past marker. Schematically, this can be represented as follows:

1. Future: (*u*) (*ma*)/*f^{ad}* + “to be: Aorist” + *f(aj)* + ADJ/PTC
2. Past: PST marker + *u* + “to be: Perfective” + *f(aj)* + ADJ/PTC

Here, again, in Taghzut and Bunsar, either conjugated or nonconjugated adjectives and participles can be negated in this way. In Zerqet, more commonly, only nonconjugated adjectives and participles are negated in this way. In Taghzut and Bunsar, the preverbal negator *u* is typically absent when there is another preverbal particle, such as the future marker (*ma*)/*f^{ad}*. In Zerqet, the preverbal negator *u* is usually present also in this case. It can also be noted that the initial (optional) element *ma* in the future marker *maf* underwent a reanalysis as the first negator *ma* (from

Table 24.12 Negation of Berber adjectives with reference to the future and past

	Taghzut	Bunsar (Tamadit)	Zerqet (Ikherruden)
1)	(u) f ^ˆ að iʒi-γ f məq:r ^c -aγ	(u) f ^ˆ að ij:i-γ f məq:ur ^c -aγ	(u) f ^ˆ að ili-γ f məq:ur ^c -aγ) ^a
2)	indi w l:i-γ f məq:r ^c -aγ	(i)ʒa w ḍɜ:i-γ f məq:ur ^c -aγ	(i)ʒa w ḍɜ:i-γ f məq:ur ^c -aγ) ^b

^aMore common: *nək u f^ˆað ili-γ f məq:ur^c* (1S NEG1 FUT^ˆIRR be.A-1S NEG2 big:MS) “I will not be big/old” (with a nonconjugated adjective)

^bMore common: *iẓa w ḍɜ:i-γ f məq:ur^c* (PST NEG1 be.P-1S NEG2 big:MS) “I was not big/old” (with a nonconjugated adjective)

Table 24.13 Negation of Arabic adjectives and participles in the future and past

ADJ <i>məzjan</i> “good”		PTC <i>sakən</i> “dwelling”	
Taghzut	Bunsar	Taghzut	Bunsar
“I will not be good”		“I will not be dwelling”	
(u) f ^ˆ að iʒi-γ f məzjan-iγ	(u) f ^ˆ að ij:i-γ f məzjan-iγ	(u) f ^ˆ að iʒi-γ f sakn-iγ	(u) f ^ˆ að ij:i-γ f sakn-iγ
“I was not good”		“I was not dwelling”	
indi w l:i-γ f məzjan-iγ	(i)ʒa w ḍɜ:i-γ f məzjan-iγ	indi w l:i-γ f sakn-iγ	(i)ʒa w ḍɜ:i-γ f sakn-iγ

Arabic), which made the Berber negator *u* redundant in this case. Nevertheless, the combination of *u* + *ma* is possible in Taghzut, both in the reference to the future (where *ma* can be analyzed as part of the future marker *maɸ*) and in the reference to the present, where *ma* cannot be analyzed as part of the future marker any more, as in *u ma sakn-iγ f* “I am not dwelling.” Table 24.12 illustrates the negated examples with the conjugated forms of the adjective *məq:(u)r^c* “big/old”: 1) “I will not be big/old”; 2) “I was not big/old.” Table 24.13 provides examples with the negated 1S conjugated forms of the Arabic adjective *məzjan* “good” and participle *sakən* “dwelling” (Taghzut/Bunsar).

Relative Forms of Berber Adjectives, Arabic Adjectives, and Participles

Introduction

Berber languages have a special verbal form traditionally known as “the Berber participle.” Here, this is referred to as “the relative form,” to avoid confusion with participles borrowed from Arabic. This form is used in subject-relative clauses and has the scheme *i-STEM-n*, e.g., *i-çeɸm-ən* “the one who entered” (pan-Senhaja). Senhaja varieties differ in the form of the relative marker: *a* in Ketama and Taghzut, and *n(:)a* in the rest of Senhaja (Seddat, Hmed, Bunsar, and Zerqet). (It is not always possible to distinguish between the anaphoric clitic *n:a* and the relative marker *n(:)a*,

and which one of the two can be absent. In Zerqet, between the relative marker *na* and the relative form starting on *i-*, the epenthetic *ð* can be inserted. In this chapter, the relative marker is written as *nað* in this case. Also, in Zerqet, the relative marker can be a single *n*.) Relative forms can function as modifiers, e.g.,

- (17) (a) *argaz* *a* *j-sʰah-ən* *i-d:a* =*d* (K)
 man RM RF-be.strong.PVF-RF 3MS-go.PFV =VC
 (b) *arjaz* *nað* *i-sʰih-ən* *i-wsa* =*d* (Z)
 man RM RF-be.strong.PVF-RF 3MS-come.PFV =VC
 “A man who is strong came.”

Adjectives and participles can occur in contexts of relativization. In Senhaja, only Berber adjectives have morphologically derived relative forms. Relative forms of adjectives function as modifiers. There are dialectal differences in the way the relative form of Berber adjectives is built. Arabic adjectives and participles can occur in relative constructions but have no morphologically derived relative forms.

Berber Adjectives in Relative Constructions

In Ketama, Taghzut, and Hmed, the relative form of Berber adjectives is built by the suffix *-n* alone and lacks the prefix *i-*. In Bunsar/Zerqet, the relative form of adjectives has the regular scheme as found with verbs, *i-* . . . *-n*. In Seddat, where cognate lexemes are obligatorily conjugated with regular PNG markers, the prefix is optional: (*i-*) . . . *-n*. As with verbs, the relative form of adjectives normally follows the relative marker *a* or *na*. The relative form is unmarked for gender and number. However, in Taghzut, the relative marker preceding the relative form is *a* with a singular referent, and *i* with a plural referent. This is different from relative constructions with verbs, where the same marker *a* is used in all contexts. Alternatively, one could posit that the relative marker is *a* in the singular (followed by the prefixless relative form), and that there is no relative marker in the plural (while *i* is the prefix of the relative form, as with regular verbs). In either case, there is a difference between the singular and plural forms, whether it is expressed in the relative marker or in the prefix.

Besides the relative form, there exists an alternative relative construction that can be used in relative clauses. In Ketama and Taghzut, its scheme is *a* (relative marker) + *jl:a* (a petrified 3MS Perfective form of the verb “to be”), followed by nonrelative forms of adjectives. A variant with a relative form of the verb “to be” (*je-lla-n*) also exists but is not frequent. A similar construction is also found in Hmed, where the form of the verb “to be” is optional. When it is present, it can be frozen (*jl:a*), relative (*jəl:an*), or conjugated. In the rest of Senhaja (Seddat, Bunsar, and Zerqet), an equivalent of this construction with the verb “to be” in the Perfective is possible but is only used with the reference to the past. The schemes and examples follow. The relative markers are included in the schemes and examples (Table 24.14).

Table 24.14 Berber adjectives in contexts of relativization**Scheme 1. Relative form of Berber adjectives in Senhaja**

Ketama	Seddat	Hmed	Bunsar/Zerqet	Taghzut	
a X-n	na (j-)X-n	na X-n	na j-X-n ~n(að) i-X-n	SG	a X-n
SG=PL	SG=PL	SG=PL	SG=PL	PL	i X-n

Example: relative form of *məq:(u)rʰ* ‘big’: ‘who is/are big’

Ketama	Seddat	Hmed	Bunsar/Zerqet	Taghzut	
a məq:rʰu-n	na (j-)məq:rʰ-ən	na məq:urʰ-ən	na j-məq:urʰ-ən, n(að) i-məq:urʰ-ən	SG	a məq:rʰ-ən
SG=PL	SG=PL	SG=PL	SG=PL	PL	i məq:rʰ-ən

Scheme 2. Berber adjectives in relative constructions (Ketama, Taghzut, Hmed)

	Ketama/Taghzut	Hmed
MS	a jl:a X	na (jl:a~jə-l:a~jə-l:a-n) X
FS	a jl:a X-θ	na (jl:a~jə-l:a-n~θə-l:a) X-θ
PL	a jl:a X-n	na (jl:a~jə-l:a-n~l:a-n) X-n

Example: relative construction with *məq:(u)rʰ* ‘big’: ‘who is/are big’

	Ketama	Taghzut	Hmed
MS	a jl:a məq:urʰ	a jl:a məq:urʰ	na (jl:a~jə-l:a~jə-l:a-n) məq:urʰ
FS	a jl:a məq:rʰu-θ	a jl:a məq:rʰ-əθ	na (jl:a~jə-l:a-n~θə-l:a) məq:urʰ-əθ
PL	a jl:a məq:rʰu-n	a jl:a məq:rʰ-ən	na (jl:a~jə-l:a-n~l:a-n) məq:urʰ-ən

Examples follow:

- (18) (a) *lʕið* *a* *məz:i-n* (K/T)
lʕið *na* *mətf:iç-ən* (H)
 festival RM small-RF
- (b) *lʕið* *n(að)* *i-məzʰ:ij-ən* (Z)
 festival RM RF-small-RF
 “Eid which is small” (Lesser Eid, Eid al-Fitr). Cf. *lʕið a məq:rʰu-n* (K),
lʕið a məq:rʰ-ən (T), *lʕið na məq:urʰ-ən* (H), *lʕið na j-məq:urʰ-ən*
 (B/Z), and *lʕið n i-məq:urʰ-ən* (Z) lit. “Eid which is big” (Bigger Eid,
 Eid al-Adha, Festival of the Sacrifice)
- (19) (a) *ibəf:afən* *na* *mətf:iç-ən* (H)
 sheep:PL RM small-RF
- (b) *ibəf:afən* *nna* *j-məzʰ:ij-ən* (Z)
 sheep:PL RM RF-small-RF
 “sheep that are small/young”

Arabic Adjectives and Participles in Relative Constructions

There are no morphologically derived relative forms of Arabic adjectives and participles in Senhaja. This is different in Ghomara, where both Berber and Arabic adjectives, as well as borrowed participles, have a relative form (Mourigh 2016). In Bunsar and Zerqet, seemingly, relative forms of Arabic adjectives are possible, with the same scheme as with the regular verbs and Berber adjectives (*i-X-n*). However, such relative forms are found only with those lexemes that have been incorporated into the regular verb class in these varieties, while those lexemes that function exclusively as adjectives have no relative forms. Therefore, such relative forms are derived from verbs rather than from adjectives.

When Arabic adjectives or participles occur in relative constructions, they simply follow the relative marker (*a*, *na*), with the frozen form of the verb “to be” in Ketama/Taghzut, optionally present (and optionally conjugated) in Hmed (cf. scheme 2 with the Berber adjectives discussed above). In Zerqet, there are dialectal differences regarding relative constructions with Arabic adjectives and participles. Such constructions are possible in the dialect of Bunjel but are rare in the dialect of Ikheruden, where relative forms of cognate verbs are used instead. In Bunjel, relative forms of verbs are used alongside relative constructions based on cognate adjectives.

Table 24.15 presents the scheme and examples of Arabic adjectives and participles in relative constructions. For Zerqet, the Bunjel dialect is shown. Arabic adjectives can have external or internal plurals, so the suffix *-in* in the scheme does not apply to all examples. For example, the adjective *məzjan* “good” has an external plural (by means of the suffix *-in* alone), while the adjective

Table 24.15 Arabic adjectives and participles in relative constructions

Scheme			
	Ketama/Taghzut	Hmed	Seddat/Zerqet
MS	a jl:a X	na (jl:a~jə-l:a~jə-l:a-n) X	na X
FS	a jl:a X-a	na (jl:a~jə-l:a-n~θə-l:a) X-a	na X-a
PL	a jl:a X(PL)(-in)	na (jl:a~jə-l:a-n~l:a-n) X(PL)(-in)	na X(PL)-in
Examples:			
Relative constructions with the adjective <i>məzjan</i> “good”: “who is/are good”			
	Ketama/Taghzut	Hmed	Seddat/Zerqet
MS	a jl:a məzjan	na (jl:a~jə-l:a~jə-l:a-n) məzjan	na məzjan
FS	a jl:a məzjan-a	na (jl:a~jə-l:a-n~θə-l:a) məzjan-a	na məzjan-a
PL	a jl:a məzjan-in	na (jl:a~jə-l:a-n~l:a-n) məzjan-in	na məzjan-in
Relative constructions with the participle <i>sakən</i> “dwelling”: “who is/are dwelling”			
	Ketama/Taghzut	Hmed	Seddat/Zerqet
MS	a jl:a sakən	na (jl:a~jə-l:a~jə-l:a-n) sakən	na sakən
FS	a jl:a sakn-a	na (jl:a~jə-l:a-n~θə-l:a) sakn-a	na sakn-a
PL	a jl:a sakn-in	na (jl:a~jə-l:a-n~l:a-n) sakn-in	na sakn-in

tʰwil “tall” forms plural by apophony (the plural stem is *tʰwal*) in addition to the suffix (hence: *tʰwal-in*).

Examples follow:

- (20) (a) *argaz a jl:a tʰwil, βaβa* (K/T)
 man RM be:3MS:PFV tall:MS father
 (b) *arjaz na (jl:a) tʰwil, (δ) βaβa* (H)
 man RM be:3MS:PFV tall:MS PRED father
 (c) *arjaz na tʰwil, δ baba* (Z-Bunjel)
 man RM tall:MS PRED father
 “The man who is tall, is my father.”

Compare the relative forms (invariable for gender and number) of the related verb *tʰwal* (Ketama/Hmed), *tʰwil* (Zerqet) “to be(come) tall”:

Ketama	Hmed	Zerqet
a j-tʰwal-ən	na j-tʰwal-ən	na j-tʰwil-ən

In Ketama and Hmed, there is a difference in meaning between the following two examples: example (21a) with an adjective in the relative construction describes a state, while example (21b) with a verbal relative form describes a result:

- (21) (a) *a jl:a tʰwil* (K)
 RM be:3MS:PFV tall:MS
 “who is tall”
 (b) *a j-tʰwal-ən* (K)
 RM RF-be.tall.PFV-RF
 “who became tall”

In Zerqet, the difference between the two constructions is dialectal: example (21a) with an adjectival relative construction is used in Bunjel, while example (21b) with a verbal relative form is common to Bunjel and Ikherruden. The meaning remains the same. To express a result, the verb “to become” in combination with an adjective is used in Zerqet:

- (22) (a) *na tʰwil* (Zerqet: Bunjel)
 RM tall:MS
 (b) *na j-tʰwil-ən* (Zerqet: Bunjel/Ikherruden)
 RM RF-be.tall.PFV-RF
 “who is tall”

The following example illustrates the use of the relative construction with an Arabic adjective *mes:us* “insipid” in a sentence. In the Hmed variety, the word *məs:us* can be used either as a Berber adjective or an Arabic adjective.

(23)	<i>nek:i</i>	<i>u</i>	<i>s:-ay</i>	<i>fi</i>	<i>ataj</i>	<i>a</i>	<i>jl:a</i>	<i>mes:us</i> (K)
	1S	NEG	drink.IPF-1S	NEG	tea	RM	be:3MS:PFV	insipid
	<i>nek</i>	<i>u</i>	<i>s:-ay</i>	<i>f</i>	<i>ataj</i>	<i>na</i>	<i>mes:us</i>	(H/Z-Bunjel)
	1S	NEG	drink.IPF-1S	NEG	tea	RM	insipid	

“I do not drink tea that is insipid.”

Summary and Conclusion

In Senhaja, adjectives and participles constitute distinct morphological classes. There are native Berber and borrowed Arabic adjectives, differing in morphological marking, but fulfilling the same function. Borrowed Arabic participles share some features with adjectives: They can modify the head noun; they distinguish the same number of forms (masculine singular, feminine singular, and plural); and Arabic adjectives with external plurals and participles share the same suffixes to mark the gender and number. In Ketama and Hmed, adjectives are never conjugated, while in Seddat, lexemes cognate to Berber adjectives are obligatorily conjugated with the regular PNG affixes (prefixes and suffixes) and are distinguished from verbs only by the lack of aspectual distinctions. For Seddat, such words can be analyzed as defective verbs. Some other varieties (Taghzut, Bunsar, and Zerqet) have both conjugated and nonconjugated Berber adjectives. Conjugation may be carried out by suffixes only (Taghzut, Bunsar, the Bunjel dialect of Zerqet) or can be regular (the Ikheruden dialect of Zerqet). In Taghzut and Bunsar, Arabic adjectives and participles can also be conjugated. This conjugation is carried out by suffixes only.

When adjectives and participles (conjugated or not) appear on their own, they refer to the present, or no specific time reference is made. The reference to the future and to the past is expressed by means of the auxiliary verb “to be.” There are different strategies to negate adjectives and participles: the negation can be achieved by the negated form of the verb “to be”, by a single negator *mafi*, or by a bipartite negation. In Zerqet, the strategy depends on whether the adjective is conjugated or not.

In subject relative clauses, a special form (a relative form) can be used. In Senhaja, only Berber adjectives have morphologically derived relative forms. Arabic adjectives and participles can occur in relativization contexts but have no morphologically derived relative forms.

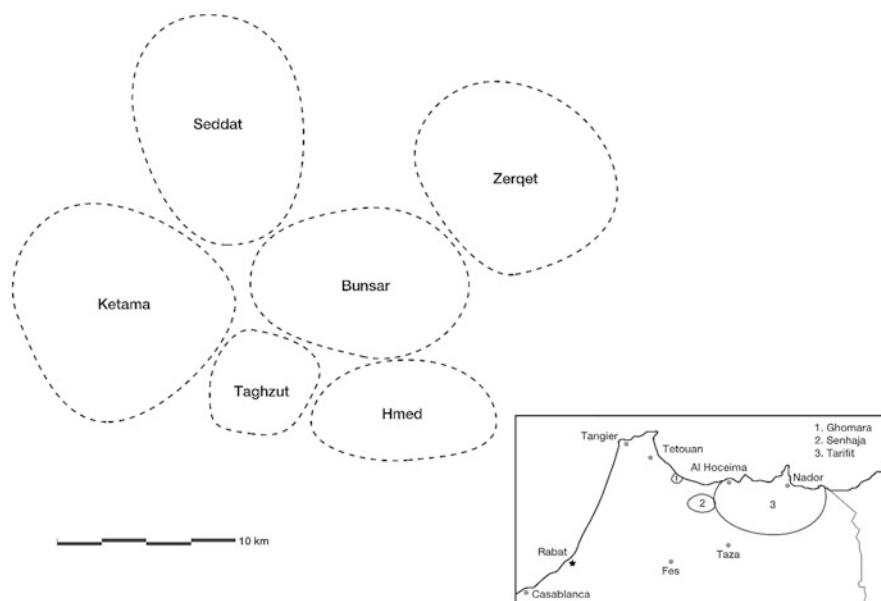
The following features of adjectives in Senhaja are important to underline:

1. Adjectives form a separate morphological class, with distinct markers.
2. (Unconjugated) Adjectives have no prefixes in Senhaja.
3. When adjectives are used predicatively, no copula (predicative particle) is used in any Senhaja variety; by contrast, when nouns are used predicatively, a copula can be used in Hmed and is required in Zerqet.
4. Adjectives are conjugated in parts of Senhaja, making them quasi-verbs.

5. When conjugated, adjectives either have a special marking carried out by suffixes only (Taghzut, Bunsar, and parts of Zerqet), or a regular marking (Seddat, parts of Zerqet).
6. Berber adjectives tend to have a particular pattern (long C2), e.g., *məq:ur^ε* (pan-Senhaja) “big,” *məz:i~məz^ε:ij* (most varieties) “small,” *mətf:iç* (Hmed) “small,” and *mes:us* (can function as an Arabic or a Berber adjective in Hmed) “insipid.”

The suffixal conjugation of adjectives and participles recalls the stative conjugation found in many Berber varieties. However, conjugated adjectives in Senhaja are not stative verbs, because they lack the aspectual distinctions. Conjugated adjectives can be regarded as an intermediate category between adjectives and verbs. There is a continuum between nonverbs and verbs: Adjectives differ from verbs in that they mark only gender and number (and not person) and in that they have no aspectual distinctions. Conjugated adjectives mark the person in addition to number and gender, which makes them similar to verbs, while they lack aspectual distinctions.

A map of Senhaja varieties discussed in this chapter



References

- Chaker S (1985) Adjectif (qualificatif). In: Encyclopédie berbère, vol II, pp 129–136, Edisud.
<https://journals.openedition.org/encyclopedieberbere/857>. <https://doi.org/10.4000/encyclopedieberbere.857>
- El Hannouche J (2008) Ghomara Berber. A brief grammatical survey. MA thesis, Universiteit Leiden

- Galand L (1990) Du nom au verbe d'état: le témoignage du berbère. In: Mukarovskij HG (ed) *Proceedings of the fifth international Hamito-Semitic congress: Hamito-Semitic, Berber, Chadic*, vol I. Afro-Pub, Wien, pp 123–138
- Galand L (2002) *Études de linguistique berbère*. Peeters, Leuven-Paris
- Gutova E (2021) *Senhaja Berber Varieties: Phonology, Morphology, and Morphosyntax*. PhD thesis. Sorbonne
- Gutova E (2022) Caractéristiques principales du senhaja (berbère du nord-ouest du Maroc). *Études et Documents Berbères* 48:65–112. <https://doi.org/10.3917/edb.048.0065>
- Haut Commissariat au Plan (HCP). Royaume du Maroc (2014) Recensement général de la population et de l'habitat (RGPH). [Morocco census 2014]. http://www.hcp.ma/downloads/RGPH-2014_t17441.html. Consulted 01 July 2017
- Ibáñez E (1959) *Diccionario español-senhayí (dialecto bereber de Senhaya de Sraïr)*. Instituto de Estudios Africanos, Madrid
- Kossmann M (2009) La flexion du prétérit d'état en berbère: questions de morphologie comparée. In: Mettouchi A, Chaker S, Philippson G (eds) *Études de phonétique et linguistique berbères: Hommage à Naïma Louali (1961–2005)*. Peeters, Paris, pp 147–167
- Kossmann M (2017) La place du parler des Senhaja de Sraïr dans la dialectologie berbère. In: Vicente Á, Naciri-Azzouz A, Caubet D (eds) *La région du Nord-Ouest marocain. Parlers et pratiques sociales et culturelles* 12. Presses de la Universidad de Zaragoza, Zaragoza, pp 93–105
- Lafkioui M (2007) *Atlas linguistique des variétés berbères du Rif*. Rüdiger Köppe Verlag, Köln
- Lafkioui M (2009) Les indices personnels verbaux des variétés berbères du Rif. In: Bouhjar A, Souifi H (eds) *L'Amazighe dans l'Orientale et le Nord du Maroc: variations et convergences*. Editions IRCAM, Maroc, pp 109–118
- Mourigh K (2016) *A grammar of Ghomara Berber (North-West Morocco)*. Rüdiger Köppe Verlag, Köln
- Renisio A (1932) *Études sur les dialectes berbères des Beni Iznassen, du Rif et des Senhaja de Sraïr*. Grammaire, textes et lexique, vol 22. Ernest Leroux, Paris
- Taine-Cheikh C (2003) L'adjectif et la conjugaison suffixale en berbère zénaga. In: Lentin J, Roth A (eds) *Mélanges David Cohen*. Maisonneuve et Larose, Paris, pp 661–674
- Taine-Cheikh C (2009) À propos de l'expression de l'état en zénaga. Apophonie et sous-catégorisation verbale en berbère et en arabe. In: Baldi S (ed) *Studi Magribini (Nuova Serie)*, vol VII (International Afro-Asiatic Congress. 11th–13th September 2008), Naples "L'Orientale", pp 95–109
- Taine-Cheikh C (2014) Qualification and comparison in Berber. The verb-noun distinction and its fluctuations. *STUF-Lang Typol Universals* 67(1):63–79

The Use of the Bare Aorist in Figuig Berber (Eastern Morocco)

26

Maarten Kossmann

Contents

Introduction	548
Basics of the Figuig Berber System of Mood and Aspect	548
The Aorist	550
Framed Aorists	552
After <i>mi</i> “when” and <i>daj mi</i> “every time that”	553
After <i>i</i> “when”	554
Other Aspectual Framings	555
Chained Aorists	555
Introduction	555
Bare Aorists in Chains Without a Clear Initial Verb	557
What Does the Chained Aorist Express?	559
The Chained Aorist in Orders	559
The Chained Aorist Continuing a Perfective	562
Conclusion	567
Transcription and Abbreviations	567
References	567

Abstract

This article studies the use of the Bare Aorist in the Berber variety of Figuig, an oasis at the border of Morocco and Algeria. Like elsewhere in Berber, the Bare Aorist is used as a form that is neutral as to aspect, whose interpretation is provided by context. It is shown that the major types of occurrence that were established for other varieties (Mauri 2017) are also represented in Figuig Bare Aorists. Framed Aorists are found in main clauses that are preceded by a temporal/aspectual clause that provides the framing for its aspectual interpretation. With Chained Aorists, the framing is provided by verbs in preceding main clauses. It is shown, however, that Chained Aorists are subtly different in Figuig

M. Kossmann (✉)
LUCL, Leiden University, Leiden, The Netherlands
e-mail: m.g.kossmann@hum.leidenuniv.nl

from the way they have been described for other varieties, as they mainly occur to express continuity at points where such a continuity is less expected.

This article studies the use of one specific morphological form, the Aorist, as a way to create clause coherence (Mauri 2017) in the variety of Berber spoken in Figuig. Figuig is an oasis on the Moroccan side of the Moroccan–Algerian border, with about 15,000 inhabitants. The study is based on recordings made in the early 1990s. The recordings were put into written text by the author with the help of a number of native speakers, and later on collated from the original, and once more retranscribed. The corpus mainly consists of traditional narratives; in addition to this, one longer recording was made, in which two older female speakers discuss traditional weddings, and the disappointing turn wedding practices have taken presently. The corpus consists of about 4.5 h of fast-paced speech, totaling about 38,000 words, produced by ten different speakers.

Figuig Berber has been the subject of a large number of studies, including a grammatical description (Kossmann 1997), a profound study of verbal morphology (Saa 2010), a study of dialect variation (Ben-Abbas 2003), and several dictionaries and word lists, most importantly Benamara (2013) and Yeou (2022). Moreover, Hassane Benamara has edited four volumes with texts (Benamara 2011, 2016, 2020; Benamara and Abbou 2021), and the present author has published a series of articles on different pragmatic and stylistic questions, based on the same corpus that is used in the present article (Kossmann 2015a, b, c, 2016; cf. also Kossmann 2000). Except for a few pages in Kossmann (1997: 348–351), no study of the uses of the Aorist has been undertaken so far, for this variety.

Introduction

Basics of the Figuig Berber System of Mood and Aspect

Berber languages express most of their aspectual and modal information by the choice of a specific form of the verb stem. In the Figuig Berber system, verbs have three different stems in affirmative sentences and two in negative sentences, the latter always combined with a negative particle. These stems will be called MAN (mood/aspect/negation), following Heath (1999: 154). In Figuig Berber, there are maximally five different stem forms, as illustrated below by means of the verb *atəf* “to enter”; there is, however, large-scale homonymy between certain MAN stems. The forms in the Table 25.1 are 3SG.M. (marked by *j-* ~ *i-*).

The Perfective is mainly used for dynamic punctual events in the past and for stative situations, irrespective of time; the Imperfective is used for a wide range of imperfective meanings, including iterative, habitual, and progressive. As for the Aorist, see the next section.

In addition to MAN stem choice, a number of other elements are used to convey aspectual and modal distinctions. One of these is found in the large majority of

Table 25.1 Overview of the aspectual stems in Figuig Berber

	Affirmative		Negative
Aorist	<i>j-atəf</i>		
Perfective	<i>j-utəf</i>	Negative perfective	<i>ul j-utif (faj)</i>
Imperfective	<i>i-t:atəf</i>	Negative imperfective	<i>ul i-t:itəf (faj)</i>

Berber varieties: the particle *ad*, which is used to express that the event has not yet been realized. This may imply a simple reference to future events about which one is relatively sure, but it can also be more modal in its usage. Moreover, it is regularly used to express habitual events, without any modal implications. In most Berber languages, *ad* can be combined, both with the Aorist (for punctual events), and, in modal and future contexts, with the Imperfective (for iterative or habitual events).

Many Berber languages also have supplementary devices, which differ from language to language. In Figuig Berber, there are three important modal/aspectual constructions in addition to MAN stem choice and *ad*-constructions:

sad + Aorist / Imperfective: – This is basically used in the same contexts as *ad* + Aorist / Imperfective. While its exact use remains to be analyzed in full, it seems that *sad* is more insistent in its pragmatic implication, suggesting a higher degree of certainty when the event is likely to occur, and suggesting more insistence when the verb is, rather, an adhortation. The label “Future” in Kossmann (1997) may be slightly misleading.

ili (“to be”) + Perfective / Imperfective – This is an auxiliary construction that focuses on the event as a situation, rather than on the event as a development. It is very common with Perfectives conveying states, and, in combination with the Imperfective, in progressives. These are, however, not the only contexts where the construction is found, and a precise analysis of its uses is still lacking.

q:im (“to sit”) + Imperfective. This is an auxiliary construction expressing an event of longer duration.

The constructions with *ili* and *q:im* can be used, both with full subject-agreement marking, and in truncated forms (Kossmann 1997: 367).

Morphologically, there is large-scale homonymy between the Perfective and the Aorist, depending on the formal structure of the verb. To provide some illustration, I counted the first 150 affirmative Perfective and Aorist verb forms in the story *N:ajər Bugr^əəm* as told by <A>. In this count, I excluded the high-frequency verb *ini* “to say,” auxiliary uses of *ili* “to be” and *q:im* “to stay,” as well as Imperatives. I found that 120 have a form that is, morphologically, ambiguous between Perfective and Aorist, while 30 mark the opposition morphologically – that is, only 20% of the cases. As there are many contexts where only the Aorist, or only the Perfective, is allowed, it is often possible to decide on the basis of context which aspect is meant, that is, which aspectual form would surface if one would substitute the verb with a verb showing the opposition. Thus, the Perfective cannot co-occur with the modal particle *ad*, while following certain auxiliaries, while in subordination the Aorist without *ad* is impossible. However, once a certain context allows for both aspects, it is impossible to differentiate between the two. In the following, homonymous forms, disambiguated by their syntactic context (e.g., after *ad*), are marked according to the

intended aspect, while in situations where the context provides no clue they are marked $A = P$.

The Aorist

The Aorist occurs in three main types of constructions:

- (a) In combination with the modal particles *ad* and, in Figuig, *sad*. In addition, in Figuig and in other languages, it is found after the subordinator *ald ~ and* “until”
- (b) As an Imperative, receiving special Imperative inflectional markers
- (c) Without a modal particle and with normal inflection

The third construction will be called “Bare Aorist” here, following Belkadi (2013), Galand (2016), and Mauri (2017). Within “Bare Aorist” constructions, Mauri (2017) distinguishes three sub-types:

1. Free Aorist – a relatively rare construction, where an Aorist with normal inflection (i.e., which is not an Imperative) is used as an initial verb in a main clause. In Figuig, this construction almost always expresses adhortative or injunctive mood (see also Kossmann 1997: 348).
2. Apodosis Aorist – a construction where the main clause has the Aorist when following a temporal/aspectual subordinate clause. As non-clausal types of temporal framing may also lead to the use of the Bare Aorist, I prefer to call this a Framed Aorist.
3. Chained Aorist – a construction where the Aorist is used in a non-initial main clause, continuing the aspect of a preceding verb. Chained Aorists are always affirmative in meaning, and therefore never follow a negative initial verb. One may note that, while this seems to be the case in most Berber languages, Adagh Tuareg may be an exception. Cf. the following sentence, where a Bare Aorist follows a negative verb: *wɛr tək:ən məd:ən ɛq:əlɛn-d* “nobody would pass (NI) (there) and come back (A)” (Ag Erless 1999: 26, line 2, transcription adapted). In this sentence, the Negative Imperfective *wɛr tək:ən* “they would not pass” is continued by the Aorist *ɛq:əlɛn-d* “they come back,” which receives its habitual interpretation from the preceding verb.

The last two categories are only found in non-initial position in an utterance. Both can take different aspectual interpretations. Thus, depending on the nature of the subordinate clause, the main clause Aorist, under the Framed condition, can be equivalent to a Perfective or an Imperfective (referring to a habitual, for instance), while the Chained Aorist can take most aspectual values available in the language, depending on the initial verb in the chain. As the Framed Aorist and the Chained Aorist both imply the presence of a preceding element providing the aspectual/temporal frame, Galand (2016; and elsewhere) is no doubt right in considering them two syntactic sub-types of the same larger phenomenon.

The use of the Framed Aorist seems to be general among those varieties that regularly use the Bare Aorist in the non-initial position; thus, it is attested in Tashelhiyt (Galand 2010: 230), Ayt Atta Central Moroccan Berber (Mauri 2017: 201); Kabyle (Manseri 1999: 52); Tuareg (Prasse 2009: 267–268; Kossmann 2011: 155), Ghadames (Kossmann 2013a: 172), and Zenaga (Taine-Cheikh 2016). There are dialectal differences as to whether conditional subordination can be followed by a Bare Aorist too. This is possible in Ahaggar Tuareg and in Zenaga, but not, for instance, in Figuig and Kabyle.

There are some major differences between languages when it comes to the distribution of the Chained Aorist (condition 3). In most languages of Morocco, the Chained Aorist can be used following any type of aspect/mood – Perfective, Imperfective, Imperative, *ad* + Aorist. In some languages, such as Tashelhiyt, this includes stative predicates (Galand 2010: 229). In Tuareg, a similar range of uses is attested, but the Chained Aorist is only rarely used as a continuation of a Perfective (Kossmann 2011: 154). In Ghadames, no examples of Chained Aorists continuing a Perfective verb were found (Kossmann 2013a: 171), while in Tarifit the use of the Chained Aorist is restricted to the continuation of Imperatives. As far as we know, the Chained Aorist is absent in Siwa and in Libyan varieties, other than Ghadames (Galand 2010: 231). The situation in Kabyle is complicated. The Bare Aorist seems to be used mainly as a continuation of the Imperative, but also appears – albeit rarely – in narrative contexts. In the latter function, its interpretation does not necessarily depend on the verb preceding immediately. According to Reesink (1979: 240), “[t]out ceci montre qu’en kabyle il ne s’agit pas d’une forme verbale liée ou enchaînée comme c’est le cas pour le chleuh, le touareg et le chaoui [. . .]”. While some authors have interpreted this as a sign that the Chained Aorist is a receding form (Chaker 1983: 227; Galand 2010: 230–231), Manseri (1999: 44) considers it “assez vivant” but it is not entirely clear from her formulation whether this includes injunctive usage of the Aorist.

There seem to be important dialectal differences in the frequency of the use of Chained Aorist forms. They are, for instance, very common in Tashelhiyt and, in the proper contexts, in Ghadames, where long series of Bare Aorists are easy to find. In such languages, it seems that a break in the series of Aorists is triggered, either by a change in aspectual reference, or indicative of a major break in the consecution of events. Put otherwise, the Chained Aorist merely indicates that the events are linked, and only a change in aspect, or another kind of strong break in the consecution of events, force one to abandon the series of Bare Aorists. Things may be different in other varieties. Thus, in Ghomara Chained Aorists, mainly (although not exclusively), appear as the last element of a chain of events (Mourigh 2015: 340). In this case, the Chained Aorist would rather function as a marker of a forthcoming break. The situation is rendered opaque, however, by the fact that in Ghomara (like in Figuig) most verbs use the same form for the Aorist and the Perfective. In Ghomara, the verb *d:u* ‘to go’ makes the difference between the two aspects. As this verb is often the closure of a series of events, it may be difficult to decide whether it is the final form after a series of Perfectives, or part of a longer series of chained Aorists.

Prasse (2009: 258) has observed that in Tuareg the Chained Aorist is incompatible with topicalization of an argument. The same is true in Figuig: there is no

example in the corpus, where the clause containing the Chained Aorist is preceded by a topicalized subject or object, which, in view of the high frequency of topicalization constructions in the language (Kossmann 2016), is probably not accidental. I have not been able to find references to this phenomenon for other Berber languages.

Most discussions concerning the Chained Aorist have concentrated on its structural interpretation. Thus, for Galand (1977, 2002 [1987], 2003, 2010, 2016; and elsewhere) the Chained Aorist is a major argument in favor of his general analysis of the Aorist as an aspectless form, which receives its aspectual interpretation from the linguistic context. From a different perspective, Mauri (2017) endeavors to integrate Chained Aorist constructions into typological models of clause chaining, arguing that it represents the rare type of initial-medial clause chaining, and formulates its use as a form of (syntactic) deranking – that is, a reduced predicate which extracts parts of its meaning from higher ranked predicates. The exact patterns of usage of the Chained Aorist have sparked less interest, the most explicit being the following, often cited, statement by Bentolila on Ayt Seghrushen (Central Morocco) (for a similar statement on Ahaggar Tuareg, see Prasse 2009: 258):

en effet, quand on emploie en SV₂ [that is, the non-initial verbal syntagma, MK] un “non-enchaîné” on rapporte les actions isolément : elles forment comme des îlots indépendants, sans relation, sans point de contact. Au contraire, quand on emploie un “enchaîné” le procès s’enchaîne au précédent, sans solution de continuité – ce qui peut produire un effet de rapidité – et se présente comme sa conséquence logique. (Bentolila 1981:153–154)

While the element of rapidity has been challenged by Galand (2010: 232), the idea that the Chained Aorist is especially used for close continuation seems to be generally accepted.

In the present article, I will present the Framed and the Chained Aorist as used by speakers from Figuig when telling traditional stories, and, in one case, giving a description of wedding traditions. My focus will be on the pragmatic meaning conveyed by such use, as opposed to non-Aorist forms, rather than on the general characterization of the Bare Aorist.

It will be shown that the Framed Aorist is quite common, and similar in use as described for other languages. The Chained Aorist, while frequent, seems to have a far more restricted use than suggested by analyses such as that given by Bentolila (1981).

Framed Aorists

Figuig Berber often employs the Bare Aorist in a main clause preceded by a temporal/aspectual subordinate clause. This can be interpreted as a kind of temporal/aspectual framing: the time/aspect expressed in the subordinate clause provides an expectation as to what aspect would be appropriate in the main clause. When this expectation comes true, the Bare Aorist can be used.

This use of the Bare Aorist is only possible when the subordinate clause precedes the main clause. Different from Zenaga and Tuareg, the Bare Aorist is not acceptable in the apodosis of conditional subordinations.

Like in many other Berber languages (see Kossmann 2013b: 351ff., for an overview), the temporal/aspectual interpretation of a subordinate clause is mainly conveyed by the choice of subordinator and not by the choice of the verb form in the subordinate clause. Thus, for example, *mi* is used in clauses expressing a habitual or iterative event, while *i* is used in clauses expressing a single anterior event. Even though *mi* clauses, thus, semantically, correspond to imperfective events, while *i* clauses correspond to perfective events, the verb in the subordinate clause almost always takes the Perfective in both cases.

The Bare Aorist is found, both after *mi* clauses and after *i* clauses. After *mi* clauses, it expresses a habitual, durative, or iterative event, while after *i* clauses it expresses a single, perfective event.

After *mi* “when” and *daj mi* “every time that”

We will first study subordinations introduced by *mi* “when” and *daj mi* “every time that”. The particle *daj* (dialectally also *ndaj*) is the normal Figuig word for ‘just, only’. The same particle is attested in Ouargla *daj* (Delheire 1987: 61). There is no relationship, it seems, to the Middle Atlas preverbal particle *daj*, which is used in narrative contexts and is always followed by an Aorist (Bisson 1920: 110). Subordinations with *mi* and *daj mi* refer to events that are repetitive or that are general truths. The main clause, therefore, refers to an event, which in non-framed constructions would be expressed by an Imperfective. In such main clauses framed by a (*daj*) *mi* subordination, the Bare Aorist often appears.

- [1]. *akəd mi ɣ:-tus an: n tməɾʰ:ut sik: ijam taf m:is kulfi igər:əɣ daj d ifuf:a. təq:arʰ dəx. . .*

“And every time this woman would come back (P) from drawing water, she would find (A) her son full of vomit (A = P), all saliva. She would say (I) then. . .” <A>

- [2]. *iwa daj mik: rʰahən n fa n umfan, daj mi ɣ:-tus tini-jas. . .*

“Well, every time they would go to some place (A = P), every time she would come (P), she would say (A). . .”

- [3]. *daj mi das-ɣ:-k:sən itəf*

“Every time they would cut off for him a part (of the meat) (A = P), he would eat (A)” <C>

The speaker has the choice between a Bare Aorist and an explicitly marked durative, habitual, or iterative verb, that is, the Imperfective, *ad* + Aorist (in its habitual interpretation), or the *q:im* + Imperfective durative construction.

- [4]. *ja an: n təmzʰa daj mi sat trʰah l frʰa n umfan it:izar-it aju n hdidwan*

“Well, this ogress, every time she intended to go (*sad* + A) to some place, Hdidwan went before her (I)” <Z>

- [5]. *daj mi q̣:-usən ə: war:a n:əs at təybər*

“Every time her children would come (P), she would hide (*ad* + A)” <C>

- [6]. *daj mi d ə: mi təf:əy d ṣ̣:qaa q̣:im tʃajarənt-tət tqəj:arin*

“Every time she would go out (A = P) into the alley, the (other) girls would scold her (*q̣:im* + I)” <A>

These aspectually explicit verb forms seem to be roughly equivalent to the Bare Aorist in this context. Thus, exactly the same state of affairs that is expressed by *q̣:im* + I, in example [6] (from the Sarsara story, Kossmann 2000: 105–115), is expressed by a Bare Aorist in a version of the same tale told by another story teller:

- [7]. *iwa daj mi təf:əy l ṣ̣:qaaq inint-as tqəj:arin*

“So, every time she would go out (A = P) into the alley, the (other) girls would say to her (A)” <C>

The use of the Bare Aorist in this construction is found with most narrators.

After *i* “when”

The subordinator *i* is used for anterior temporal subordinations. The main clause, therefore, expresses a punctual event, mostly with past reference. In non-framed contexts, the Perfective would be used for such a situation. In main clauses framed by *i* the Bare Aorist is often encountered.

- [8] context: a boy has been eaten by an eagle, and the eagle is coerced to give him back in good shape.

i t-id-ijru izbəṛ̣ idʒal; in:a-jasən i ləl n:əs

“When he (= the eagle) had thrown him back (P) in good shape (A = P), he (= the boy) took an oath (A) and said (P) to his family” <A>

- [9] *i das-iyərəs uk: un: n tʒərzərt im:atər kulʃi d aməl:al*

“When he had slaughtered (A = P) the gazelle, he saw (A) that everything was white” <A>

- [10]. *ik: rəʃlən ja ik: iq:əs ja r:ʃəl nsən, nət:a jas-q̣: ij: iman n:əs t taməṭ:ut*

“When they married (A = P), when their wedding started (A = P), he came (A) and pretended (A) he was a woman”

The same construction is found when *i* is preceded by *al ~ an* “until”.

- [11]. *təs:əwa-q̣: s:aləf n:əs jawkan an i dis juləj təm:atr-i d argaz, tən:a-jas*

“She let down (A = P) her hair and then (until) when he had climbed up (P) by it, she saw (A) that he was a man, and she said (P)” <A>

Among the older female story tellers, who use the Bare Aorist regularly in a sequence of perfective events (see below), the Bare Aorist is all but obligatory after a subordination with *i*, the exceptions being a few cases with *ini* “to say”. The other story tellers mostly use the Perfective, e.g.,

- [12]. *i tənt-itfu, tus-q: jəm:atsənt*

“When he had eaten them (P), their mother came (P)” <D>

- [13]. *i tədʰsʰu jawka im:utər taymās: ə: tadalt jisi-t:*

“When she laughed (P) he saw (P) her green tooth and took her (A = P)” <O>

Other Aspectual Framings

The framing by means of a subordinate clause is by far the most common type of context for a Framed Aorist. However, the same usage is sometimes found in contexts that have a similar effect, but that are not formally temporal/aspectual subordinate clauses (cf. for similar constructions in Central Moroccan Berber, Mauri 2015: 337). This is especially the case in distributional contexts, e.g.,

- [14]. *iwa mərʰ:a mərʰ:a jawka jini-jas-q: ja*

“Well, from time to time he would say to him (A)” <O>

- [15]. *iwa daj aɣəʰ:uf irʰaħən, ə: iyərs-as əju n ɣmərʰ, aɣəʰ:uf irʰaħən inəy bab n:əs*

“Well, any palanquin that would go out (A = P), this Omar would slaughter them (A=P), any palanquin that would go out (A = P), he would kill (A) its master”

- [16]. *si q:-if:əy suk: jəz:ar wi s: m:a ilqan jini-jas*

“From the moment that he left the butcher, everyone who would meet him (A = P) said to him (A)” <M>

Chained Aorists

Introduction

In Figuig, Chained Aorists are found following initial verbs in the Perfective, in the Imperfective, in the non-realized (*s*)*ad* + Aorist constructions, and with an Imperative. I did not find unambiguous cases where the Aorist continues a verb in auxiliary constructions with *ili* “to be” or *q:im* “to keep on”.

Chained Aorists do not have to be close to each other, or to the initial verb, and many elements may come between them as long as they are not verbs in another aspect. Moreover, as will be clear from many of the following examples, there is no need for subject continuity. From this point of view, the link between the initial verb and the following Aorist is rather weak.

Chained Aorists only continue affirmative statements. I have not found any unambiguous cases where a stative verb appears in a Chained Aorist construction.

Note that the use of consecutive stative Aorists has been described for Tashelhiyt (Galand 2010: 229) and Ayt Atta (Mauri 2015) but is unknown in Tuareg (Prasse 2009: 259), and Kabyle (Galand 2010: 230, citing Manseri 1999). The following examples illustrate the use of the Chained Aorist after different initial verb forms.

PERFECTIVE:

- [17]. *jawkan tuy-i təmsi. jas-ǧ: n jəm:as in:a-jas. . .* “Then he got (P) fever (lit. fever passed at him). He went (A) to his mother and said (P). . .” <A>
- [18]. *i tud'a təsxəf, i ǧ:-usən ajətməs usən-ǧ: usən-ǧ: walu u dasən-ǧ:-təj:i ləmʕaf, usən-ǧ: walu. afən-t: tud'a*
 “When she had fallen down (P) unconscious (P), when her brothers came (P), they came (P), came (P), to no avail, she had not made (NP) them their food, they came (P) to no avail. They found (A) that she had fallen down (P)” <O>

IMPERFECTIVE:

- [19]. *mukud təǧ:in zʕ:rʕudʕəj:a s:irdən-t:, awjən-tət-ǧ: i f:warijat*
 “Then they (habitually) grind (I) carrots, wash them (A = P), and bring (A) them in bags” <B; description>
- [20]. *mukud tʃər:ən ǧ:unajt tʃədʕ.*
 “Then everybody would pick (I) and grind (A)” <B; description>

AD + AORIST AND SAD + AORIST

- [21]. *a ǧ:-tas ja s lyiwan, at təs:bingər ja, tas-ǧ: tadən ayənsu n:əs, tas-ǧ: ja s lyiwan*
 “She (the bride) comes (*ad* + A) with singing, she puts on her veil (*ad* + A), she comes (A) and covers (A) her face, she comes (A) with singing” <B; description>
- [22]. *tan a fək-t:əf tətʃ-iji*
 “Look, she will eat you (*ad* + A) and eat me (A)” <A>
- [23]. *irʕa a ǧ:-nawj aʕəq:a nawəj-ǧ: Jaja Ambrʕuka a kidəm trʕaħ*
 “Wait (IMPT), we will bring (*ad* + A) a necklace, we will bring (A) Yaya Ambrouka, so that she will go (*ad* + A) with you” <C>
- [24]. *sat trʕaħəd at tafəd pʕ:atnəx itʕ:əsʕ, sa das-tqəs:əd tmart n:əs, thəs:nəd tqəs:d-as af:arən n:əs təj:d-as tisijt z:atəs*
 “You must go (*sad* + A) and you will find (*ad* + A) our father asleep (P), you must cut (*sad* + A) his beard, cut his hair (A = P), cut (A = P) his nails, and put (A) a mirror in front of him” <A> (As the wider context and interpretation do not allow for a Perfective, one can be sure that *thəs:nəd* and *tqəs:d-as* are Bare Aorists.)

IMPERATIVE

Imperatives are normally in the Aorist. What is called here a Chained Aorist after an Imperative is the use of a non-Imperative second person Aorist verb after an initial Imperative.

- [25]. *j:ət azʕʕal, ad iħma qbala tadərm-as dis fus n:əs a dawəm-təfdʕəħ*
 “Prepare (IMPT) soup, let it get very hot (*ad* + A) and put (A) her hand into it
 so that she will reveal it to you (*ad* + A)” <E>

In none of the abovementioned contexts, the use of an Aorist is obligatory. Its use is relatively rare with *ad* + Aorist and with *sad* + Aorist. Cases like the following example ([26]) are much more frequent in the corpus than are cases with Chained Aorists:

- [26]. *am:ən: jif: n jumajən nix təlt ij:am iwa ad j:ən dəx miq:ən id-w. . . id-sʕ:luħat.*
a s:-ij: abəʕda urgaz d aməzwar, ad j:ən amənsi qaʕ i lʕaʔilt n:əs, ləl n:əs, d at xalis, d
at ʕəm:is ja d sʕ:ləħ.
mi dəx jif: n təlt ij:am nix rʕbəʕ ij:am ad ij: dəx pʕ:as ə: ajən: . . . ad ij: dəx pʕ:as n təsləjt
dəx aju n umənsi am:u.
a dasən-jini dəx ik: at tiq:art: (. . .)
ad ʕəj:dʕən dəx i ləl n:ə: . . . i ləl nsən d at xalitsən d at ʕəm:itsən,
iwa ad q:imən miq:ən ja i tiq:art nsən.
taməʔ:ut, twafunt at təq:im i tiq:art n urgaz n:əs.
 “Like this, two or three days later, people prepare (*ad* + A) ritual meals.
 First the husband does it (*ad* + A), they make (*ad* + A) dinner for all of his family, his
 relatives, the family of his maternal uncle, the family of his paternal uncle, as a
 ritual meal.
 Some three or four days later, the father makes (*ad* + A) this. . . the father of the bride
 makes (*ad* + A) this dinner like this.
 He says (*ad* + A) to the people of the house: (. . .)
 They invite (*ad* + A) the relatives of. . . their relatives, and the family of their maternal
 uncle and their paternal uncle,
 and the people stay (*ad* + A) in their house.
 The woman, the girl, stays (*ad* + A) in the house of her husband” <B; description>

This long passage about the weeks following the wedding has the construction with *ad* + Aorist (in its habitual usage) throughout. While there is no question about the consecution of the described events, there is not a single Chained Aorist in this fragment. With Perfectives and Imperatives, the Chained Aorist seems to be more common, while our corpus has too few instances of Imperfectives followed by another verb with Imperfective aspectual meaning to make any statements about the frequency of the Bare Aorist vs. the Imperfective in these contexts.

Bare Aorists in Chains Without a Clear Initial Verb

While the immediate consecution of the initial verb and the Aorist verb is, by far, the most common situation, there are a number of examples where the aspectual matrix of the Aorist is provided by a verb that is further away from it, or where it is not obviously continuing a preceding form (cf. Mauri 2015: 336, for similar constructions in Central Moroccan Berber), e.g.,

- [27]. *imsək:ar-q: akidəs iwd'a g^w: un n tasa. təq:jəm təls'əg dis l'əfəjt. tali iq:ar: "(...)" iwa ir'əy din iwa tr'əj:əh xfas*
 "He got up angry (A = P) with her and fell (A = P) into the pitfall. There was (q:im + P) fire there. Then he started to say (I) "(...)". Well, he got burned (A) there and she (had) got rid (A = P) of him" <F>

In this fragment, if *ir'əy* "he burned" were a continuation of the preceding verb (*iq:ar*), it should be interpreted as an Imperfective. This is not entirely inconceivable, but the more probable interpretation is that *ir'əy*, just like *tr'əj:əh xfas* "she (had) got rid of him," constitutes a concluding statement for the scene which would normally be expressed by a Perfective.

- [28]. *iwa k:rən z'wan l ləxla, isin qa kulfi, t:an luqid. yərsən i tfunas:, səlxən-t:, irzan maj z:əg ala s:ənwən ul ufin. iwa q:im t:əq:lən, m:atrən amz'a uk: d'r'ar'...*
 "So, they went (k:ər + A = P) to the desert, they took (A = P) everything with them, they forgot (A = P) the matches. They slaughtered (A = P) the cow, they skinned (A = P) it, they sought (A = P) with what they would be able to cook it (ad + A), but they couldn't find it (NP). So, they started to look (q:im + I) and saw (A) an ogre on the mountain. ..." <C>

In example [28], the immediate context of the Aorist *m:atrən* "they saw" is the durative construction *q:im t:əq:lən* "they looked for a long time". There is no doubt that *m:atrən* is meant to convey a punctual event without any specific duration. One could argue that the auxiliary verb *q:im* (here in a truncated form) is formally a Perfective, and therefore, the aspectual continuity is not broken, but this seems to be strained.

- [29]. *k:urənt k:urənt and awd'ənt jif: ə: n sənt n trəgwin. if: n tərğa n uyi d jif: n tərğa... həfak n ngət'r'an. ... tək:ur tək:ur tən:a-jas jawkan awd'ənt, tən:a-jas...*
 "They walked (I) and they walked (I) until they arrived (ald + A) at two irrigation canals. One canal with buttermilk, and one canal – excusez le mot – with tar... She walked (I) and walked (I), she said (P), then they arrived (A) and she said (P)... " <D>

Example [29] is complicated, as the storyteller corrects himself. In the second sentence, he starts with *tən:a-jas* "she said," and then rephrases it by means of *jawkan awd'ənt tən:a-jas* "and then they arrived (A) and she said (P)," followed by the speech verb and the direct speech. In this constellation, it is highly improbable that the Aorist *awd'ənt* would be a continuation of the preceding Perfective *tən:a-jas*, which it corrects. In fact, it is not clear how the Aorist *awd'ənt* in this sentence should be analyzed as continuing any preceding main verb form, as all other verbs are (durative) Imperfectives.

- [30]. *iwa idʒən n was: dəx, tən:a-jas: “ax-am at təs:irdəd u... ulman.” təq:im təs:irid ulman təq:im tfət:sr-in i... i sʕ:dʕuħ. im:atər-tət-q: idʒən n ə: n m:is n uʒəl: id, ik:ər iwtu-q: idʒən n uxəlɣal ʕlaħsab dʕa... dʕarʕ n:əs*
 “So, one day, she said (P) to her: ‘here you are, you should wash (*ad* + A) the warp threads’. So, she started to wash the warp threads (*q:im* + I), she started to spread them out (*q:im* + I) on the roof terrace. A son of the king saw her (A), and he made (P) an anklet in the size of her foot” <O>

Like in the earlier examples, the immediately preceding verbs are durative in aspect (*q:im* + I), while the event described by the Aorist (*im:atər-tət-q:* “he saw her”) is a punctual, perfective event. One remarks that the verb immediately following the Aorist, *iwtu-q:* “he hit, he fabricated,” which describes a direct consequence of the seeing, has the Perfective form, combined with the semantically bleached verb *k:ər* “rise,” which is often used to convey the beginning of a new series of events.

What Does the Chained Aorist Express?

The optionality of the use of the Chained Aorist confronts us with the following question: how does its use relate to that of fully specified aspectual marking in similar contexts? For this problem, we will compare the use of Chained Aorists in the two contexts for which it is amply attested in our corpus – situations where the initial verb is an Imperative and situations where it is a Perfective.

The Chained Aorist in Orders

As mentioned above, Chained Aorist forms of Imperatives are characterized by the fact that they take normal second person marking, rather than Imperative forms. Normal second person Aorist forms are not used in initial position.

When several orders are given, three different constructions are possible. In the first construction, Imperatives are used throughout, e.g.,

- [31]. *rʕuħ awj-iq: taqəq:it:, sʕiwdʕ-iq: amrʕəlʕ:a*
 “Go (IMPT), bring me (IMPT) meat, bring me (IMPT) a lot” <C>

In the second construction, the Imperative is followed by *ad* + Aorist. Different from the Bare Aorist, *ad* + Aorist can also be used in (polite) orders in initial position, so this is essentially a series of two different types of order. It is not possible to start the series of orders by *ad* + Aorist, and continue with an Imperative.

- [32]. *rʕuħ a xəfnəx təzʕwid*
 “Go and get away from us” <O>

In the third construction, the Imperative is followed by a Chained Aorist (see ex. [25] above).

In longer series of orders, it is possible (but not obligatory) to start with several Imperatives, and then continue with Chained Aorists, e.g.,

- [33]. *a r^ʕuħ l ud^ʕr^ʕar^ʕ, s:əħma din jif: n tyunt s:əħma-t: mliħ, təj:əd s waḏ:wəs abərda, taljəd s uzən:a n ud^ʕr^ʕar^ʕ tək:zəd s uqəlqul*

“Go (IMPT) to the mountain, heat up (IMPT) a stone, make it (IMPT) very hot, put (A) sand underneath, go up (A) to the top of the mountain and jump down (A = P) with your head (downwards)” <D>

There are no strict rules governing the choice between these constructions, but there are strong tendencies. The IMPT – *ad* A construction is mainly found following the Imperatives *(a)r^ʕwaħ* “come!”, *r^ʕuħ* “go!”, and *k:ər* “get up”. All three are special verbs, albeit for different reasons. *(a)r^ʕwaħ* only exists in the Imperative – The reason to consider it a verb form is that it has different forms with singular and plural addressees and functions as a suppletive Imperative to *as-q:* “to come”; *r^ʕuħ* has been borrowed in its Imperative form from Maghribian Arabic, while all other verb forms are based on a different Arabic stem form, *r^ʕaħ* (see Kossmann 2013b: 267). It is the only verb in the language that has a special Imperative stem that is different from the Aorist. Finally, *k:ər* is often used as a verb indicating the beginning of a new event, which does not necessarily mean any focus on the action of getting up. *K:ər* often occurs in this way in narrative sections, but also plays a role in Imperatives, e.g.,

- [34]. *k:ər at tḡəd^ʕr^ʕəd!*

“Come (IMPT) eat lunch (*ad* + A)!” (lit. get up and you shall eat lunch) <E>

The IMPT – *ad* A construction is particularly common with *(a)r^ʕwaħ*: over half of the attestations of the construction are with this verb. There are only very few cases in the corpus where *(a)r^ʕwaħ* is followed by another Imperative, and there are no attestations in the corpus where it is followed by a Chained Aorist. It is not clear why this is the case, and why *(a)r^ʕwaħ* “come!” is much more common in the IMPT – *ad* A construction, than the semantically similar Imperative verb *r^ʕuħ* “go!”.

As regards the choice between the IMPT – IMPT and the IMPT – A construction, there does not seem to be such a strong lexical determination. The main difference between these two constructions lies in the proximity of the two orders in the syntactic string. IMPT – IMPT constructions are almost exclusively found when the second order immediately follows the initial Imperative, with the possibility of having some discourse markers in between, e.g.,

- [35]. *aw-awim, aju n təjdət: ḡd^ʕat-tət əj:t-as ndaj ajk: ixəs*

“Here you are, this dog, guard it (IMPT) and give to it (IMPT) just what it wants” <F>

- [36]. *wa, awəj-t-id*

“Go down (IMPT) and bring it (IMPT)” <M>

The construction is common when the first verb is the Imperative *rʕuħ* “go!”, which, being an intransitive verb, is of course often constructed with a following order without any element intervening, e.g.,

[37]. *rʕuħ p:ab-t-id*

“Go (IMPT) and carry him here (IMPT) on your back!” <D>

Once a series of Imperatives is started, it is possible to have successive orders in the Imperative, with elements intervening, e.g.,

[38]. *rʕuħ isi fəm: . . . aɣjul, isi dis iqlilən, j: iman n:əm təl:id tʕ:əɣ:nəd*

“Go (IMPT) and take (IMPT) . . . a donkey, put (IMPT) pitchers on it, and pretend (IMPT) you are passing by (*ili* + I)” <E>

In example [38], the first two Imperatives follow each other immediately (*rʕuħ* and the first *isi*), but the other Imperatives (the second instance of *isi* and *j:*) are separated by intervening elements.

The construction with a Chained Aorist, on the other hand, is almost exclusively used when there are elements separating the initial Imperative from the later order(s), e.g.,

[39]. *ɣərs uk: ufriɣ, tisid-ɣ: aqəlqul, d jiləm, təs:əwit-tən i təzɣawt d aməzwar*

“Slaughter (IMPT) a ram, take (A = P) its head and hide, and let them first down (A=P) (into the well) in a basket” <A>

The number of elements in between the two orders can be considerable, as in example [40]:

[40]. *uɣt-as qaɣ an: n nəksəwt (n:iɣ-awəm) uk: un: n urgaz, uk: un: n urgaz nə: ɣ: ɣərwəm il:a jus, Məx:adʕ. iwa tufm-as an: n twafunt nwəm (. . .)*

“Give (IMPT) all the clothes (I tell you) to that man, to that man eh that has come to you, Churner. Then give to him (A = P) your daughter (. . .)” <A>

All in all, the choice between the IMPT – IMPT construction and the IMPT – A construction can, to a very large degree, be predicted from the absence, or presence, of intervening elements. It should be noted that the correlation is not perfect, as shown by the following two exceptions – from the same scene, told by the same storyteller – where two Imperatives are separated by an intervening direct object:

[41]. *a jəm:a awj-iɣ: afuħ n wudi. jr-i-t-id i təħrirt*

“Mother, bring me (IMPT) some butter and throw it (IMPT) into the soup” <Z>

[42]. *awj-iɣ: a jəm:a afuħ n wudi jr-iɣ: dis i ə: . . . i təħrirt inu*

“Bring me (IMPT) some butter, mother, and throw it (IMPT) into it e: . . . into my soup” <Z>

The Chained Aorist Continuing a Perfective

The use of the Chained Aorist when continuing a Perfective is linked to a number of factors. In the first place, there are clear differences among storytellers. In order to study this, a full-scale comparison of the ratio between Bare Aorists and Perfectives was undertaken (Of course, only those forms that formally distinguish between Aorist and Perfective were included, and only those Aorists that continue a Perfective event were counted.). While among Aorists only the Chained forms were included, any Perfective was counted, as long as it was not part of a subordinated clause. Therefore, it is no wonder that all storytellers opt more for Perfectives than they do for Bare Aorists. Moreover, as we will see later, the extremely frequent verb “to say” has a strong tendency to appear in the Perfective, thus further skewing the results toward the Perfective. As far as I can see, these caveats have a bearing on the numbers, but not on their relative frequency among speakers. I do not have the impression that some speakers would have fewer consecutive situations, or fewer instances of “to say” than others.

The results are highly revealing. Among the older female storytellers, , <C>, and <F> had ratios between one Bare Aorist to 4 Perfectives to one Bare Aorist to 6 Perfectives. Only the older female storyteller <A> had a lower ratio, with one Bare Aorist to 12 Perfectives; this may be due to her specific stylistic choices in using this construction (see below). The other story tellers used the Bare Aorist much less frequently in this context. The stories by <O>, for example, only have one Bare Aorist to 40 Perfectives. Among the male storytellers, Bare Aorists were even rarer. The reason behind these differences seems to lie, to a large extent, in story telling experience. The highest ratios are found among the most experienced storytellers. Only <A>, a blind lady in her fifties (in 1990), well-known in her community as a storyteller, has a slightly lower ratio, but still has a score over three times higher than <O>. Most of the other narrators had much less experience in telling stories; or had long not done so when recorded. Thus <O>, when recorded, did not live in Figuig anymore and, even though she knew the stories very well and delivered them with great delight, her performance was less confident than that of, for example, <A>, as shown by many corrections and abundant stopgaps. The same is true for most of the male performers, who would not typically be telling this type of stories (see for a similar effect of narrative experience on another grammatical feature, Kossmann 2015b: 220).

It has often been claimed that in some languages the use of the Chained Aorist in perfective contexts is receding. This is suggested, for example, by Galand (2010: 230–231) for Kabyle, and by Bououd (2014: 122) for Ayt Sadden Central Moroccan Berber. The Figuig situation could be interpreted in the same way, as the oldest performers have a higher incidence of Bare Aorists than younger performers. As the Chained Aorist in perfective contexts seems to be a form that is strongly correlated to some types of narrative style, I prefer to interpret the differences as reflecting the degree of knowledge of stylistic conventions. The fact that the two young men in the corpus hardly used any Bare Aorists in Perfective consecution would, thus, not so much, be caused immediately by an age-gradient in the variation, but rather due to a

lack of experience in the genre. Of course, one interpretation goes hand in hand with the other. When a certain construction is strongly related to a specific genre, and the genre in question gets less commonly performed – as is no doubt the case with traditional narratives in Morocco – it will become more and more marginal.

In addition to stylistic preferences (and maybe competence), the use of the Chained Aorist in Perfective consecution is also restricted in the case of one specific verb. The verb *ini* “to say” is among the most frequent words in the corpus, as it is the normal introduction to direct speech. Much of the direct speech is, of course, part of a series of events, and, moreover, dialogue, itself, often includes series of direct speech events. Because of this, one would expect *ini* to appear regularly in the Chained Aorist. This is, in fact, the case when repetitive events are described. Thus, in the following passage, a repeating dialogue is reported, starting by the (suppletive) Imperfective form *q:ar^c* “to say,” followed by a Chained Aorist:

[43]. *daj mi d ə: mi təf:əy d s^c:qaq q:im tʃajarənt-tət tqaj:arin:*

- *r^cuh xəfnax a Sarsara-wəḍ:ər-səbʃa!*

tas-ḍ: n jəm:as til: təq:ar^c-as:

- *a jəm:a, mar diḍ:-q:ar^cənt Sarsara-wəḍ:ər-səbʃa?*

tini-jas:

- *a jəl:i hən:a a ndaj.*

“Every time she would go out (A = P) into the alley, the girls would scold her (I):

- Go away from us, Sarsara-that-made-seven-stray!

She would go (A) to her mother crying (A = P = I). She would say to her (I):

- Mother, why do they call me Sarsara-that-made-seven-stray?;

and she would say (A):

- Well, my daughter, that’s the way it is” <A>

However, when the dialogue is a single event that is not repeated, the Chained Aorist is hardly ever used. This is illustrated by the following dialogue from a story by <F>, a storyteller who, otherwise, uses the Chained Aorist regularly (example [44]). In order to make the example easier to process for the reader, false starts have been substituted by [***]. None of them appear in a context that is relevant to the issue at stake.

[44]. *iwa təq:jəm til: al axtən kud tas-ḍ: lal:as tən:a-jas:*

- *maj ʃəm juɣən til:əd?*

tən:a-jas:

- *a maj diḍ:-in:a m:is n buhəp:u [***], a maj diḍ:-in:a izmər, a maj diḍ:-in:a.*

tən:a-jas:

- *ini-jas: isəm:ən:-ək u ikəp:ər-ək [***] ndbəh-lik. u nakül ləhm-ək u nəgʃəd ʃla idurt-ək.*

*iwa təq:jəm ja al ajətfə tərgəb [***] ja tən:a-jas:*

- *ja jəl:is n nan:a.*

tən:a-jas:

- *ja m:is n buhəp:u, isəm:ən:-ək, u ikəp:ər-ək, u nakül ləhm-ək u nəgʃəd ʃla idurt-ək.*

“So, she started to cry (*q:jəm* + I). Until a certain moment, then her mistress came (A) and said (P):

- What’s the matter with you that you are crying?

She said (P):

- This is what the son of Buheppu said to me. This is what the lamb said, this is what he said.

She said (P):

- Tell him: He will fatten you and make you big, and I will slaughter you and eat your meat and sit on your skin. [formulaic sentence in Arabic]

So, she waited (A = P) until the next day and she looked down (A = P) (from her window) and he (the lamb) said (P) to her:

- O, daughter of my mistress!

And she said (P):

- He will fatten you and make you big, and I will eat your meat and sit on your skin”

<F>

In order to compare the use of the Chained Aorist as opposed to the use of the Perfective under similar conditions, it is therefore necessary to focus on the older female story tellers, and also to leave the verb “to say” out of the equation.

As mentioned before, Bentolila, in his analysis of the use of Ayt Seghrushen consecutive forms considers them as marking continuity (especially in fast succession) and logical succession (Bentolila 1981: 153–154). Applying Bentolila’s analysis to our Figui data, one would expect the Chained Aorist (when continuing a Perfective) to be especially common in simple chains of events, and to be less frequently found when there is a conceptual break in the chain, for example, when something unforeseen happens. Moreover, even though Bentolila does not mention this, one could imagine that a serious pause in the performance would lead to a break in the chain, and, thus, would not be likely to be followed by a Chained Aorist (cf. also Galand 2016: 452).

These predictions are not confirmed in the Figui corpus, cf. the following passage:

[45]. *iwa ald itf qa tafunas:, ip:r-in ik: iləm, jas:əs xəfsən, irʿaħ irəḍ:za-ḍ: ad ə:: mani tən ala is:nəw. [4 seconds pause] jas-ḍ: iqməf*

“Well, until he [= the ogre, MK] had eaten (*ald* + A) the entire cow, then he stuffed them (A = P) into the skin, he sewed (A) them in, and he went (A = P) searching (I) to ehh, where he would cook (A) them. [4 sec pause]. Then Iqmec came (A). . .” <C>

In this passage, the Chained Aorist *jas-ḍ:* “he came” is preceded by a pause in the story telling. Moreover, the event described by *jas-ḍ:* represents a break in the chain of events. First, we hear about the way the ogre treats the members of a certain family. After this, a brand-new protagonist is introduced, *Iqməf* (a nickname of Hedgehog), who starts a new chain of events. There is no indication that the arrival of *Iqməf* would be shortly after the ogre left; in fact, one would rather expect a certain

time lapse in between the two events. Thus, this passage, told by a confident story teller, shows a use of the Chained Aorist which is quite different from what Bentolila describes.

In fact, at least one storyteller, <A>, seems to use Chained Aorists in Perfective contexts, preferably, when there is some kind of break in the chain of events. This is illustrated in example [46] and [47].

[46]. context: A girl has been buried alive by her brother.

*iz^ɛwa iʃaf-it. jawəj-q: R^ɛap:i idʒən n ə: n. . . n nmalik isəl: i n:dir jawka
jud^ɛa-q: xʃəs*

“He went away (A=P), he ignored her (A=P). The Lord brought (A) a en king who heard the moaning and found (A=P) her” <A>

[47]. *iwa ə: r:ən-q: ja qa ləhwajəz n:əs, jaləj d lmalik*

“So, they gave back to him (A=P) all his belongings, and he ascended (A) to be the king” <A>

In example [46], a king arrives “out of the blue,” marked by a Chained Aorist. In example [47], while the focus remains on the same person, the fact that the hero of the story ascends to the status of a king is, not obvious from the preceding context, nor a necessary consequence of the general direction of this particular story.

When there is a simple chain of logically coherent consecutive events, <A> uses Perfectives rather than Chained Aorists, e.g.,

[48]. *təj:u jawkan tir^ɛəd^ɛ ayəd nət:ata ləksəwt jir^ɛd^ɛ-in jumas, tənju uk: jis təz^ɛwa*

“She did (P) (like that), she also put on (A = P) clothes that her brother had worn, she mounted (P) the horse and went away (P)” <A>

The difference between Perfective, for expected continuations, and Chained Aorists, for unexpected continuations, is also relevant in example [49].

[49]. *tədwəl-q: t tamət^ɛ:ut n:əs, iwa tur^ɛəw-q: akidəs. juʃ-as R^ɛap:i səbʃa, nət:ata an:
n nwaʃun tur^ɛəw-tən-q: təm:atr-i ja an idʒən n umul:u it. . . it^ɛr^ɛək xəʃsən ə:
lma.. lmal suk: un: n umʃan, r^ɛ:z^ɛəq, akəj:əl. usən-q: ʔəldin, q:imən tkəj:lən din
təm:atər jumas*

“She became (A=P) his wife, and she got children (P) with him. The Lord gave (A=P) her seven, she gave birth (P) to these children. Then she saw him (A). . . One day they [her brothers, MK] ran out of stuff, wealth, things to buy. They came (P) there, they started to buy (q:im + I) and she saw (A) there her brother” <A>

In this passage, the logical consequence of marrying – begetting children – is marked by the Perfective (*tur^ɛəw-q:*, *tur^ɛəw-tən-q:*). On the other hand, when a great switch takes place in the story – the main character accidentally, and unexpectedly, spots her brother, who has come from elsewhere – the Chained Aorist is used (*təm:atər*). The storyteller first mentions the fact that the main character saw her

brother, and then chooses to explain why he was away from his home, followed by a second mention of her seeing him. In both cases, the event of seeing is expressed by means of a Chained Aorist.

This use of the Chained Aorist to mark unexpected developments in the chain of events may be an idiosyncrasy of <A>'s story telling. It is not as obvious in the stories by other experienced narrators, who seem to be as likely to use the Bare Aorist with expected, as they may with unexpected events (One may note, however, that my corpora for the other experienced female story tellers are much smaller than that of <A>.).

This is not to say that cohesion is unimportant in the meaning of the Bare Aorist. In fact, one way to interpret <A>'s use of the Aorist is that it marks cohesion exactly in contexts where cohesion is not obvious. That is, it occurs when a real break in the flow of events could be envisaged, and signals that this is not the case.

The idea that the Chained Aorist is used to mark cohesion is exacerbated in narrative passages following direct speech. In traditional stories, dialogues, and other instances of direct speech, provide a specific rhythm for the narrative. Within the main chain of events, which is focused on a development from action to action (and more often than not from location to location), they constitute islands where no further action than the speech act takes place (Kossmann 2000: 63–65). As such, they are quite different from the rest of the story. In many cases, the action following a dialogue starts in a Perfective, which could be interpreted as showing a full break in the narration. It is quite possible that the preference for the Perfective with *ini* “to say” is also related to the independence of direct speech vis-à-vis the narrative parts of the story.

There are, however, many instances where the narration following a dialogue starts in a Chained Aorist. In all of these instances, the action is a direct consequence of what was said in the dialogue: as such, it would count as marking cohesion at a point where context would make one expect a full break.

[50]. *in:a-jasən*: “xsəx-t:.” *jawj-it*

“He said (P) to them: “I love her.” He married (A) her” <A>

[51]. *jən:a-jas*: “*ukan fəm:in, ndaj wik: iwjəy axtən tʃajəbd-i-t-id.*” *iny jəm:as*

“He said (P) to her: “So you vilify anybody that I marry.” (So) he killed his mother” <F>

[52]. *alaxtən dəx tədʕən lgaflət tini-jas*:

- *a dik l-gafla, la fikūm Yusfwəld ʕəm:-i?*

in:a-jas:

- *a-ni, a-ni.*

iwa jas-ɖ:

“Then a caravan came by (A = P) and she said (This is one of the rare instances where the verb *ini* “to say” takes a Bare Aorist in a sequence of perfective events.) (A):

- O caravan, isn't my cousin Yousef among you? [formulaic sentence in Arabic]

He said (P):

- Here I am, here I am. [formulaic sentence in Arabic]

So, he went there (A)” <F>

Conclusion

In this article, the main uses of the Bare Aorist in Figuig have been described. Generally speaking, these uses are similar to what has been described for other northern Berber languages. The Chained Aorist is a marker of text cohesion, which indicates that the event in question is to be seen as linked to the preceding text. The Framed Aorist functions in a similar fashion, deriving its aspectual interpretation from an expectation set by the subordinate clause.

However, in the details of its usage, the situation seems to be different in Figuig from what has been described for other Berber languages. It was shown that the use of the Bare Aorist in non-subordinate contexts is not simply a matter of strong coherence. In fact, it seems that the Bare Aorist is especially common as a marker of coherence, when coherence is less obvious from the context. This is the case of Imperatives, where the Bare Aorist hardly ever does occur immediately following the initial Imperative, and of narrative sequences, where at least one story teller uses the Bare Aorist, preferably, in contexts where coherence is not automatically implied. Thus, we may conclude that in Figuig, the Aorist is, first and foremost, a marker that signals coherence where it might be less expected.

Transcription and Abbreviations

The following abbreviations are used: A: Aorist; F: feminine; ⵓ IMPT Imperative; I: Imperfective; M: masculine; MAN: Mood-Aspect-Negation; NI: Negative Imperfective; NP: Negative Perfective; P: Perfective; sg: singular. The story tellers are identified by an anonymizing abbreviation between <> following the example. At the time of the recordings, A, B, C, F, and O were middle-aged and old women, Z a young woman, M a middle-aged man, and D and E (the latter only in dictated texts) young men. All story tellers cited in the article are from the village Zenaga, except F, who is from the village Elmaiz. In the transcription, dialectal and idiolectal differences (Ben-Abbas 2003) have not been homogenized, which leads to variation in the form of certain elements.

Following the style sheet of this volume, the transcription uses IPA symbols. It should be stressed, however, that this is not a phonetic transcription, and that many salient phonetic features are not represented. Thus /t/ is in most positions affricated ([tʃ]), and there is large-scale allophony with the vowels.

References

- Ag Erless M (1999) « Il n'y a qu'un soleil sur terre ». Contes, proverbes et devinettes des Touaregs Kel-Adagh. IREMAM, Provence
- Belkadi A (2013) Aspect and mood in Berber and the aorist issue. SOAS Working Papers in Linguistics 16:127–150
- Ben-Abbas M (2003) Variation et emprunts lexicaux. Étude sociolinguistique sur le parler amazigh de Figuig. Unpublished PhD Thesis, Fes, Université Sidi Mohamed Ben Abdellah

- Benamara H (2011) Contes berbères de Figuig (Sud-est marocain). Textes en berbère avec traductions en français. Rüdiger Köppe, Cologne
- Benamara H (2013) Dictionnaire amazighe–français. Parler de Figuig et ses régions. IRCAM, Rabat
- Benamara H (2016) Poésie berbère de Figuig (ancienne et moderne). Rüdiger Köppe, Cologne
- Benamara H (2020) Proverbes, devinettes et expressions figées berbères. Texte bilingue. L'Harmattan, Paris
- Benamara H, Abbou F (2021) Poésies de Mekki Atmane (Textes amazighes avec traduction française). IRCAM, Rabat
- Bentolila F (1981) Grammaire fonctionnelle d'un parler berbère. Aït Seghrouchen d'Oum Jeniba (Maroc). SELAF, Paris
- Bisson P (1920) Leçons de berbère tamazight. Dialecte des Aït Ndhir (Aït Nâaman). Félix Moncho. [reprint: Frontispice, 2012]
- Bououd A (2014) La grammaire de la langue amazighe: Du mot à la phrase. Accessible through <http://bououd.e-monsite.com/pages/livre/grammaire-de-la-langue-amazighe.html>
- Chaker S (1983) Un parler berbère d'Algérie (Kabylie). Syntaxe. Publications de l'Université de Provence [& Jeanne Lafitte], Provence
- Delheure J (1987) Agerraw n iwalen teggargrent–tarumit. Dictionnaire ouargli–français. SELAF, Paris
- Galand L (1977) Continuité et renouvellement d'un système verbal : Le cas du berbère. Bulletin de la Société de Linguistique de Paris 72(1):275–303
- Galand L (2002) Les emplois de l'aoriste sans particule en berbère. In: Galand L (ed) Études de linguistique berbère. Peeters, Leuven, pp 260–272
- Galand L (2003) L'aoriste berbère, l'aspect et les valeurs modales. In: Lentin J, Lonnet A (eds) Mélanges David Cohen. Maisonneuve & Larose, Paris, pp 235–246
- Galand L (2010) Regards sur le berbère. Centro Studi Camito-Semitici, Rome
- Galand L (2016) The Aorist in Berber. In: Guentchéva Z (ed) Aspectuality and temporality: descriptive and theoretical issues. John Benjamins, Amsterdam, pp 447–463
- Heath J (1999) A grammar of Koyra Chiini. Mouton, Berlin
- Kossmann M (1997) Grammaire du parler berbère de Figuig (Maroc oriental). Peeters, Leuven
- Kossmann M (2000) A study of Eastern Moroccan fairy tales. Academia Scientiarum Fennica, Helsinki
- Kossmann MG (2011) A grammar of Ayer Tuareg. Rüdiger Köppe, Cologne
- Kossmann MG (2013a) A grammatical sketch of Ghadames Berber (Libya). Rüdiger Köppe, Cologne
- Kossmann MG (2013b) The Arabic influence on Northern Berber. E.J. Brill, Leiden
- Kossmann MG (2015a) The use of the ventive marker *dd* in Figuig Berber narratives. Nordic Journal of African Studies 23(4):241–291
- Kossmann MG (2015b) Deixis in Figuig Berber narrative texts. Folia Orientalia 52:199–228
- Kossmann MG (2015c) The interplay of style, information structure and definiteness: double indirect objects in Figuig Berber narratives. Corpus 14:59–80
- Kossmann MG (2016) On word order in Figuig Berber narratives: the uses of pre- and postverbal lexical subjects. Wiener Zeitschrift für die Kunde des Morgenlandes 106:115–146
- Manseri O (1999) L'aspect en berbère : L'aoriste revisité. Awal 19:43–58
- Mauri S (2015) Formal and functional properties of grammatical aspect in Ayt Atta Tamazight. Unpublished PhD thesis, London, SOAS
- Mauri S (2017) A typological analysis of the *Chained-Aorist* construction in Ayt Atta Tamazight (Berber). Studies in Language 41(1):198–222
- Mourigh K (2015) A grammar of Ghomara Berber (North-West Morocco). Rüdiger Köppe, Cologne
- Prasse K-G (2009) Manuel de grammaire touarègue. Syntaxe. [Ms. produced in a few copies at CARGO Verlag, Schwülper]

- Reesink P (1979) Problèmes de détermination en indo-européen, principalement dans le germanique de l'ouest et dans une langue chamito-sémitique. Thèse de doctorat de 3^e cycle, Université Paris III, EPHE, IVe Section. Unpublished PhD Thesis
- Saa F (2010) Quelques aspects de la morphologie et de la phonologie d'un parler amazighe de Figuig. IRCAM (originally PhD Thesis, Paris, 1995)
- Taine-Cheikh C (2016) The Aorist in Zenaga Berber and the imperfective in two Arabic dialects. A comparative viewpoint. In: Guentchéva Z (ed) Aspectuality and temporality: descriptive and theoretical issues, John Benjamins, pp 465–501
- Yeou M (2022) Dictionnaire amazighe-français. Parler de Figuig, IRCAM, Rabat

Nataša Knežević and Amazigh Bedar

Contents

Introduction	572
Numerals and Distributivity	573
Numeral Reduplication and Pluractionality	576
Syntactic Distribution of <i>NumRed</i>	577
Participant-Distribution Is Reducible to Spatial-Distribution	580
Atomicity and Exhaustivity	586
Plurality Requirement	587
Analysis	589
Theoretical Background	589
Semantics of <i>Red</i>	590
Conclusion	593
References	594

Abstract

This chapter investigates the phenomenon of numeral reduplication (*NumRed*) in Taqbaylit, which gives rise to distributive interpretations. We argue that *NumRed* in Taqbaylit is not a universal distributive quantifier over individuals but rather a marker of event plurality. Furthermore, we show that the variety of distributive readings found with the distributive morpheme *Red(uplication)* is not due to ambiguity but rather to the vague semantics of this morpheme. Our main claims are that: (i) *Red* is an event plurality marker, (ii) distributive readings of *NumRed* arise through spatiotemporal separation of (sub)events, and (iii) participant-distributive readings should be reduced to spatial-

N. Knežević (✉)
Nantes Université, Nantes, France

A. Bedar (✉)
Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, Marseille, France

distributive readings. The aim of our investigation of *NumRed* in Taqbaylit has been to contribute insights into distributive numerals and pluractionals across languages, as well to open the door to more investigation of reduplication, distributivity, and pluractionality in the Berber languages.

Introduction

This chapter investigates numeral reduplication (*NumRed*) in Taqbaylit¹. An example with a reduplicated numeral *snaθ* “two” is given in (1)²

- | | | | | | | |
|------|--|--|--------------|--------------|------------|--------------------|
| (1) | rəfð-ənt | θəqfiʃin =ənni | snaθ | snaθ | n | təbwaʃin |
| | carry.PFV-3F.PL | CS.girl.PL =DEM | two.F | two.F | Gen | CS.box.F.PL |
| Lit. | The girls carried two two boxes . | | | | | |
| | i. Participant-distributive | “The girls carried two boxes each .” | | | | |
| | ii. Event-distributive | “The girls carried two boxes at different times/places ”. | | | | |

Reduplication is, in general, associated with distributivity across languages (Gil 1982, 1988; Balusu 2006; Bouzidi 2012; Cabredo Hofherr and Laca 2012; Kiss, et al. 2013, among others). Reduplicated numerals with distributive interpretations are also called *distributive numerals* (Gil 1982, 1988). The term *distance distributivity* (Zimmermann 2002) is used in the literature to encompass many, somewhat different, but closely related, phenomena, namely, distributive numerals (Gil 1982; Cable 2014; Knežević 2015), “anti-quantifiers” (Choe 1987; Zimmermann 2002), distributive-share markers (Gil 1995), and pluractional markers (Matthewson 2000) (In this chapter, we will use the term *distributive numerals*.).

The main difference between distributive numerals and distributive (universal) quantifiers (*mkul* in Taqbaylit, or *every* and *each* in English) is that, while the latter combine with the NP, serving as a distributive-key (the NP over which the distribution takes place) yielding so-called participant-distributive readings, the former appear to combine with the NP serving as a distributive-share (entity that is being distributed), yielding both participant-distributive and event-distributive readings, as illustrated in (1). The major issue in the literature of distance distributivity has been how to account for this apparent ambiguity. On some approaches, distributive numerals are analyzed as involving the distributive operator *each*, selecting for the distributive-key, either participants or a covert spatiotemporal argument (Choe 1987;

¹Taqbaylit, also called “Kabyle,” is a VSO language of the Berber language family (Afroasiatic phylum). The data in this chapter are from Amazigh Bedar, a native speaker of Taqbaylit of Chemini. The examples are transcribed with the International Phonetic Alphabet (IPA).

²Abbreviations: “-” = affix boundary, “=” = clitic boundary, 1, 2 & 3 = person, acc = accusative clitic (Direct object clitic), CAUS = causative, comp = complementizer, CS = construct state (Nominative case), Dem = demonstrative, Dir = Directional particle, F = feminine, FS = free state (Accusative case), Gen = genitive, IPFV = imperfective, Lit. = Literal translation, M = masculine, PFV = perfective, PL = plural, SG = singular.

Oh 2001; Zimmermann 2002; Champollion 2012). Therefore, the sentences with distributive numerals are ambiguous between participant-distributive (distribution over individuals), as illustrated with (1.i), and event-distributive readings (distribution over times/spaces), as illustrated with (1.ii).

We argue that distributive numerals (*NumRed*) in Taqbaylit cannot be analyzed as involving the distributive *each*, by showing that the distributive numerals do not involve universal quantification over individuals, but event plurality. Furthermore, we argue that sentences with distributive-numerals (*NumRed*) are not ambiguous but always involve distribution over times and/or spaces. Our main claims are that (i) *Red*(uplication) is an event plurality marker, (ii) distributive readings of *NumRed* arise through spatiotemporal separation of (sub)events, and (iii) participant-distributive readings should be reduced to spatial-distributive readings.

Although we do not contend that the distributive dependency in terms of Choe (1987) can be applied to *NumRed* in Taqbaylit, we use the terms distributive-key and distributive-share throughout the chapter to respectively refer to the NP combining with the *NumRed* as the distributive-share, and the argument that seems to be distributed over as the distributive-key (be it participant or spatiotemporal argument). We also use the term *participant-distributive reading* to describe a plural event where in each (sub)event there is an atomic participant that seems to be distributed over by the distributive-share, although we assume that participant-distribution is the subcase of the spatiotemporal distribution.

The chapter is organized as follows. We start by exposing the issue of numerals and distributivity in Section “[Numerals and Distributivity](#)”. In Section “[Numeral Reduplication and Pluractionality](#)”, we review the properties of *NumRed* in Taqbaylit by comparing them to the universal quantifiers *mkul/each*, and to pluractional markers. We show that *NumRed* always involves plurality of events that must be temporally, or spatially, separated. In some cases, only temporal distribution is allowed. In Section “[Analysis](#)”, we sketch an explicit analysis of *Red* as a pluractional marker and illustrate how it applies to our examples of sentences with *NumRed*. We conclude in Section “[Conclusion](#)”.

Numerals and Distributivity

The sentences with numerals are generally ambiguous between a collective and a distributive reading (Not only numerals are ambiguous between the collective and the distributive interpretations. This is also true of the plural NPs, for instance. In this chapter, we focus on the distributivity of numerals.). This is illustrated in the example in Taqbaylit, as shown below (The genitive preposition *n* obligatorily intervenes between a numeral and an NP).

- (2) rəfð-ənt θəqʃjʃin = ənni **snaθ** **n** **təbwaðin**
 carry.PFV-3F.PL CS.girl.PL = DEM two.F Gen CS.box.F.PL
 “The girls carried **two boxes**.”

- i. *Collective*: The girls carried two boxes **together**.
- ii. *Participant-distributive*: The girls carried two boxes **each**.

On the collective reading of (2), the girls carried two boxes together, at the same place and at the same time. On the distributive reading of (2), each girl carried two different boxes. Note that contrary to the distributive numerals (*NumRed*) in (1), numerals in (2) do not yield event-distributive readings, but only so-called participant-distributive reading.

Following Choe (1987), and subsequent approaches, distributivity is referred to as a relation between two arguments: a distributive-key (or sorting-key) and a distributive-share.

The distributive-key, which must be semantically plural, denotes the participant over which the distribution takes place, here, *θəqʃiʃin* = *ənni* ‘the girls’. The distributive-share, which must be a non-specific expression of an explicit quantity, denotes the participant (the entity) which is being distributed, here, *snaθ n təbwaʃin* ‘two boxes’. According to this, on the distributive reading of (2), each girl carried two boxes, since ‘two boxes’ are distributed over the members of the group of girls. Importantly, under the distributive reading of (2), the number of boxes depends on the number of girls, since ‘two boxes’ is multiplied by the total number of the girls participating in the event. If, for instance, there are three girls involved in the described event, the total of boxes will be six (two boxes distributed to each of the three girls).

In many languages, there are ways of disambiguating sentences with numerals by using an overt distributive marker. This can be achieved, for example, by using the universal (adnominal) distributive quantifier *mkul* in Taqbaylit, illustrated in (3a), as well as the universal quantifiers *every* and *each* (See Beghelli and Stowell (1997) for the discussion on differences between *every* and *each*. In Taqbaylit, there is only one lexical item for the (singular) universal quantifier - *mkul*. In this chapter, we will systematically use *each* in English—in parallel to *mkul* in Taqbaylit.) in English, the latter being illustrated in (3b).

- (3) a. **mkul** *θəqʃiʃθ* *θə-rfəð* *snaθ* *n* *təbwaʃin*
 each FS.girl 3F.SG-carry.PFV two.F Gen CS.box.F.PL
- b. **Each girl** carried two boxes.

Using the terminology from Choe (1987), *mkul* and *each* select as their arguments the NP ‘girl’ serving as a distributive-key (over which the distribution takes place). The NP ‘two boxes’ serves as a distributive-share, so ‘two boxes’ are distributed over the atomic members of the group of girls. Importantly, the number of boxes in the event depends on the number of girls to which ‘two boxes’ are distributed (there are two boxes, per girl, in the event).

Languages vary according to whether they have a morphological marker for the distributive-key (English), for the distributive-share (Georgian), or whether they have both distributive-key and distributive-share markers (Serbian) (cf. Gil 1988;

(4) rəfð-ənt θəqɹɪʃɪn = ənni snaθ snaθ n təbwəʒɪn
carry.PFV-3F.PL CS.girl.PL = DEM two.F two.F Gen CS.box.F.PL

Lit. The girls carried **two snag boxes**.
i. Participant-distributive “The girls carried two boxes **each**.”
ii. Event-distributive “The girls carried two boxes **at different times/places**.”

The sentence in (4) yields only distributive readings. Like the sentence with the universal quantifier *mkul/each* in (3), (4) gives rise to the so-called participant-distributive reading under which each of the girls carried two boxes. That is, “two boxes” are distributed over the atomic agent participants in the described event. Importantly, the sentence in (4), unlike the sentence in (3), also yields the event-distributive reading under which the girls carried two boxes at different times and/or places. According to Choe (1987), we can describe the event-distributive reading as distribution over a covert spatiotemporal argument. The distributive-share here is the event of carrying two boxes, which is distributed over the distributive-key – time intervals or/and locations. Note that there are two correlated differences between the distributive-key (universal) quantifiers, such as *mkul/each* in (3) and the distributive-share markers, such as *NumRed* in (4). The former syntactically combine with the NP serving as the distributive-key and yield only participant-distributive readings, while the latter syntactically combine with the NP, serving as a distributive-share, and yield two types of distributive readings: participant-distributive and event-distributive.

In the literature, the major question concerning the distributive-numerals (marking the distributive-share) is whether they are ambiguous between the participant-distributive and the event-distributive readings and how to account for this ambiguity. On the first type of accounts (Choe 1987; Zimmermann 2002), the distributive-share markers (distributive numerals), just as the distributive-key markers (distributive universal quantifiers), involve a D-operator (universal quantifier *each*) that takes as its restriction, either the participant (NP it combines with), yielding the participant-distributive reading, or a covert spatiotemporal arguments-times/places, yielding the event-distributive readings. On the second type of accounts (Cable 2014; Knežević 2015), distributive numerals (distributive-share markers) are not ambiguous and

always involve event plurality (Balusu (2006) provides an analysis of numeral reduplication in Telugu, also avoiding the ambiguity by assuming that the participant-distribution can be derived as the instance of the spatiotemporal distribution. Contrary to other operator-based approaches assuming that the distributive-key is, either a plural participant, or a spatiotemporal argument, Balusu argues that the distributive-key is always a spatiotemporal argument. However, Balusu's analysis also makes use of the distributive operator *each*, which is not the case with Cable's (2014), and Knežević's (2015) approaches.). The distributive interpretations arise through the spatiotemporal separation of the (sub)events, and the participant-distributive reading can be derived as an instance of the spatiotemporal distribution.

Following Cable (2014) and Knežević (2015), we argue that the *NumRed* in Taqbaylit is a marker of event plurality. In the next section, we examine the syntactic distribution and the semantic properties of *NumRed* by comparing them to universal quantifiers and pluractionals.

Numeral Reduplication and Pluractionality

Pluractional markers across languages often involve reduplication of different lexical categories (most typically the verb) and/or affixes on the verbs (Newman 1980, 1990; Cusic 1981; Lasersohn 1995; Landman 2000, among others). They indicate the basic meaning of multiplicity of events (actions) or repeated actions, involving, thus, multiple event times, spaces or/and participants. One of the most prominent characteristics of pluractionals is that of distributivity, i.e., "action by more than one individual, temporally iterated action, and spatially scattered action" (Lasersohn 1995: 238). This basic meaning, invoking a plurality of events, gives rise, furthermore, to the effect of the following associated meanings, listed in Cusic (1981: 74): repetitiveness, repeated occasions or events, habitual agency, distributed quality, inchoativity, distribution, cumulative result, intensity, augmentation, diminution, plurality of sites of action, duration, persistent consequences, celerativity, and continuity. Pluractional markers are classified according to three main criteria: (sub)events must have separate running times, running spaces or thematic roles (Lasersohn 1995). As to which one is chosen depends on the lexical meaning of a particular pluractional morpheme, or of the lexical category of the reduplicated item. We consider that *Red* in *NumRed* NPs in Taqbaylit is a unique pluractional morpheme responsible for all distributive readings: spatial (including participant) and temporal.

We first overview, in Section "Syntactic Distribution of *NumRed*", the syntactic distribution of *NumRed* and the kinds of participant-distributive readings depending on it. We point out that *NumRed* always yields temporal-distributive readings, unlike universal quantifiers, and like pluractionals. We also point to the fact that in some cases it also allows spatial-distributive readings, again unlike universal quantifiers, and like (some) pluractionals. Furthermore, we show in Section "Participant-

Distribution Is Reducible to Spatial-Distribution” that the so-called participant-distributive reading is reducible to the spatial-distributive reading. Then, we discuss the properties of atomicity (Knežević 2015; Knežević and Demirdache 2018) and exhaustivity (Matthewson 2000; Knežević 2015; Knežević and Demirdache 2018) in Section “**Atomicity and Exhaustivity**”, as well as the plurality requirement (Lasersohn 1995; Balusu 2006; Knežević 2015; Bedar and Allawama 2024) in Section “**Plurality Requirement**”. We show that first, unlike the distributive universal quantifiers *mkul/each*, *NumRed* does not require atomic and exhaustive distribution over individuals; second, like pluractionals, *NumRed* abides by the plurality requirement (Balusu 2006), i.e., participants provided by *NumRed*NP must be different across the (sub)events.

Syntactic Distribution of *NumRed*

Numeral reduplication (The reduplication is a very productive morphosyntactic strategy in Taqbaylit. Nearly every lexical category reduplicates (Bouzidi 2012). We hypothesize that at least the verbal reduplication also involves event plurality. However, there are some important empirical differences between the numeral reduplication and the verbal reduplication that require an in-depth examination. We leave this issue aside for further research.) (*NumRed*) in Taqbaylit appears in transitive sentences with numerals in a subject position, in an object position, or, simultaneously, in both positions. *NumRed* can also appear in transitive sentences with a morphologically singular argument (marking the singular subject or the singular object), and in intransitive sentences.

The kind of so-called participant-distributive interpretation of a sentence with *NumRed* depends on the syntactic position of the *NumRed*. Note that *NumRed* has a strict syntactic position, i.e., it cannot be split of the NP it combines with as it is the case with, for instance, the distributive universal quantifier *each* (floating *each*) in English (Beghelli and Stowell 1997). Recall that, descriptively, *NumRed* always marks the distributive-share (the entity that is being distributed).

Transitive Sentences with Two Numerals

In the examples below, we illustrate *NumRed* in an object position in (5), in a subject position in (7), and in, both a subject and an object position simultaneously in (9), and the types of readings available with them in (6), (8), and (10), respectively (The *NumRed* in transitive sentences, with one numeral and one plural NP, as illustrated in the sentence in (1), has the same semantic effect as in the sentences with two numerals (one simple and one reduplicated numeral), given in (5). We choose to present the second, because it is easier to clarify the dependence between the syntactic position of the *NumRed* and the type of the participant-distributive interpretation of a sentence.).

- (7) **0laθa** **0laθa** **n** **təqʃiŋin** rəfð-ənt snaθ n
 three three Gen CS.girl.PL carry.PFV-3F.PL two.F Gen
 təbwaʃin
 FS.box.F.PL

Lit. **Three three girls** carried two boxes.

“**Three girls** carried **each of the two boxes.**”

“**Three girls** carried two boxes **at different times/places.**”

(8) a. Participant-distributive scenario:

Last Monday, Dihya, Anya, and Kenza carried one box, while Lwiza, Damya, and Tanina carried another box at the same time.

b. Event-distributive scenarios:

Every Monday last month, (at least) three (different) girls carried two boxes. Either, three girls carried the two boxes together, or three girls carried one box, while three different girls carried another box; or, three girls carried two boxes, while three different girls carried another set of two boxes.

The sentence in (7), with *Red* combining with the subject NP “three girls” is true under the participant-distributive scenario in (8a), where, descriptively, three (different) girls (distributive-share) are distributed over the members of the set of two boxes (distributive-key), so that each of the two boxes is carried by three girls. So, when *NumRed* marks the subject (*θlaθa θlaθa n taqfiḡin* “three three girls”), the distribution is done over the object (*snaθ n tɔbwaḡin* “two boxes”).

The sentence is also true under the event-distributive scenario in (8b), where at different times/places, the groups of three girls participated in the (sub)events of carrying, cumulatively, two boxes (two boxes per subevent or two boxes in total). Again, unlike the sentences with a (simple) numeral in (2), and the universal quantifier *mkul* in (3), in the sentence with *NumRed* in (7) *the number of girls does not depend on the number of boxes* in the event, since the distribution is not (necessary) over the atomic boxes. Alike (5), *NumRed* in (7) requires at least two (sub)events involving three girls carrying boxes that are temporally and/or spatially separated. This means that there will be six girls or more, always multiples of three (nine, twelve, etc.) in the event.

NumRed in a subject and in an object positions

- | | | | | | | |
|-----|--------------------|----------|-----------------|-----------------|-------------|-------------|
| (9) | θlaθa θlaθa | n | taqfiḡin | rəfð-ənt | snaθ | snaθ |
| | three three | Gen | CS.girl.PL | carry.PFV-3F.PL | two.F | two.F |
| | n tɔbwaḡin | | | | | |
| | Gen FS.box.F.PL | | | | | |

Lit. **Three three girls** carried **two two boxes.**

“**Three girls** carried two boxes **at different times/places.**”

(10) Event-distributive scenarios:

a. *Spatial*: Last Friday, Dihya, Anya, and Kenza carried together two boxes, while Lwiza, Damya, and Tanina carried together two different boxes at the same time.

b. *Temporal*: Every Monday last month, the girls in groups of three carried the boxes, in groups of two.

In (9), *NumRed* appears in, both the subject, and the object positions. Descriptively, again, both subject and object are marked by *NumRed* as distributive-shares: they need to be distributed. Since there is no other plural argument available for the distributive-key, the distribution is over a covert spatiotemporal argument (locations and/or time intervals). This is why (9) yields only event-distributive readings in (10), and not the so-called participant-distributive reading. What is required is a multiplicity of (sub)events each of which involving the girls in threes, and the boxes in twos. These events can be happening at the same time, however, they must be spatially separated, as in (10a). Or, these events can be only temporally separated as illustrated in (10b). Parallel to (5) and (7), *NumRed* in (9) requires at least two (sub)events involving three girls and two boxes, that are temporally and/or spatially separated. This means that there will be six girls or more, always multiples of three (nine, twelve, etc.), and four boxes or more, always multiples of two (six, eight, etc.) in the event.

To resume, all transitive sentences with *NumRed* in an object position in (5), with *NumRed* in a subject position in (7), and with *NumRed* in both a subject, and an object position in (9) yield event-distributive readings (spatial and temporal). Sentences (5) and (7) also yield so-called participant distributive readings, depending on the syntactic position of *NumRed*. These are the readings under which in each subevent there are atomic participants (girls in (5); and boxes in (7)). Importantly, both atomic and non-atomic separation of participants is possible. That is, “three girls” in (5) can be separated in different subevents in groups of two and one, and not only atomically – one (atom) girl in each subevent (see scenarios in (6a) and (6b) for sentence (5)). This non-atomic participant distribution puts forward two correlated facts. First, unlike the universal distributive quantifiers *mkul/each*, *NumRed* does not distribute over (atomic) individuals. This is, at the same time, an argument against the distributive dependencies between the two arguments (in the sense of Choe 1987), which could be paraphrased as: for each girl, there are two boxes, which she carried. The availability of the non-atomic readings rules out the analysis of *NumRed* in Taqbaylit as involving the distributive operator *each*. Second, spatial separation of events is achieved via participant (atomic or non-atomic) separation.

Participant-Distribution Is Reducible to Spatial-Distribution

In the previous section, we have seen that transitive sentences with two plural arguments yield participant and event-distributive readings. Importantly, both atomic and non-atomic separation of participants is possible, which suggests that spatial separation of events is achieved through participant separation.

In this section we present two more arguments in favour of our claim that participant-distributive readings should be reduced to spatial-distributive readings. To achieve this, we present the sentences which yield only event-

distributive reading (Many pluractional markers, especially those marking the verb, yield only temporal-distributive readings. For illustration, see, for instance, Cabredo Hofherr et al. (2019) analyzing MULT verb stems (marking the multiplicity of events) in Seri, or Matthewson (2000), analyzing *pelpala7*, the reduplicated numeral ‘one’ in Lillooet Salish.) – transitive sentences with one singular argument and one *NumRed*, yielding only temporal event-distributive readings, and intransitive sentences with *NumRed*, yielding, both spatial, and temporal event-distributive readings.

Transitive Sentences with One Singular Argument

Let us look at the sentences with one singular argument: subject or object NP, and one *NumRed* NP.

Singular subject

- (11) θə-rfəð **Dihya** **snaθ** **snaθ** **n** **təbwaðin**
 3F.SG-carry.PFV Dihya two.F two.F Gen FS.box.F.PL
 Lit. **Dihya carried two two boxes.**
 # “Dihya carried two boxes **at different places/separately.**”
 “Dihya carried the boxes **two by two/two at a time.**”
- (12) Event-distributive scenarios:
- a. *Spatial:* Yesterday, Dihya carried two boxes in one hand, and two different boxes in another hand.
 - b. *Temporal:* On Mondays last month, Dihya carried two boxes up to her apartment.

As illustrated in (11), *NumRed* appears also in transitive sentences with one singular argument, here the singular subject “Dihya”. *Red* appears with the object NP *snaθ n təbwaðin* “two boxes”. The distribution of “two boxes” cannot be over the singular argument “Dihya,” and, therefore, must be over a covert spatiotemporal argument. Naturally, participant-distributive reading is not available.

Sentence (11) is felicitous under the rubric of the temporal event-distributive scenario in (12b). The (sub)events, each involving two (different boxes), have the same participant (Dihya), the same place (Dihya’s apartment), but have different running times (on Mondays).

Importantly, the sentence in (11) is *infelicitous in the spatial event-distributive scenario* in (12a), under which the (sub)events each involving two boxes have the same running time (are simultaneous), have the same participant (Dihya), but are spatially separated, i.e., in each hand, Dihya carried two boxes. This unavailability of spatial separation of events, even forced by the context in (12a), suggests that the spatial separation can only be achieved via participant separation. When participant separation is not possible, since the participant is singular (Dihya), neither participant, nor spatial distributive readings arise.

Singular object

- (13) **snaθ snaθ n təqfiʃin** ss-arð-ənt **θaxxamt = ənni**
 two.F two.F Gen CS.girl.PL CAUS-clean.PFV-3F.PL **room= DEM**

Lit. **Two two girls cleaned the room.**

“**The girls in twos** cleaned the room (at the same time).”

“The girls cleaned the room **two by two/two at a time.**”

- (14) Event-distributive scenarios:

- a. *Spatial*: Yesterday, four girls cleaned the (ball)room. Dihya and Kenza cleaned together the east part of the room, while, at the same time, Anya and Tanina cleaned together the west part of the room.
- b. *Temporal*: On Monday, Dihya and Kenza cleaned the room. On Tuesday, Anya and Tanina cleaned the same room.

In (13), the singular argument is the object NP *θaxxamt = ənni* “the room,” while the subject is marked with *NumRed*, *snaθ snaθ n təqfiʃin* “two two girls,” and, therefore, appears as the distributive-share. Naturally, the distribution cannot be over the singular object, so the participant-distributive reading does not arise.

The sentence in (13), just like the sentence in (11), is felicitous under the temporal event-distributive scenario in (14b) under which the (sub)events, each involving two different girls, have the same running space (the room), which is also the shared participant (object) of the cleaning events, but have separate running times (Monday and Tuesday).

Again, like (11), (13) *is not felicitous under the spatial event-distributive scenario* in (14a), under which the (sub)events, each involving two girls have the same participant “the room,” the same running time (are simultaneous), but have separate running spaces – west and east part of the room. Parallel to (11) and (12a), it is not possible to force the spatial separation of events in (13), by the context in (14a).

The examples with *NumRed*, and a singular subject NP in (11) and a singular object NP in (13), show that sentences with a singular argument yield only temporal event-distributive readings. The participant-distributive reading is not available since it requires a plural participant who is split into atomic members for each subevent, over which we distribute. This is not the case in (11) since the subject is the singular DP “Dihya,” nor in (13) where the object is the singular DP “the room”.

Let us turn to the unavailability of the spatial event-distributive reading. As illustrated in (12a) for (11), the only way to make the spatial separation of the subevents possible without the participant separation is imagining that Dihya carries two boxes in each hand: since there is a sole participant (Dihya), we can imagine Dihya participates in two events at the same time, each involving carrying two boxes. Nevertheless, it is not possible to use the sentence in (11), *θə-rfəð Dihya snaθ snaθ n təbwaʃin* (Dihya carried two two boxes), to describe this scenario. The same is illustrated in (14a) for (13). The only way to make spatial separation of the

(sub)events, each involving two different girls and the same participant (“the room”), at the same time, is to “distribute” “two girls” to different parts of the room. This scenario cannot, however, be described using the sentence in (13), *snaθ snaθ n təqfiʃin ss-arð-ənt θaxxamt = ənni* (*Two two girls cleaned the room*).

The question arises as to why spatial distribution is possible with (5) and (7), involving a plural subject and plural object arguments, respectively, but not with (11) and (13), involving a singular subject and a singular object argument, respectively – even when the spatial separation is made salient by the context. One possible explanation is that the only salient criterion for the spatial separation of (sub)events is via participant separation. If this is the case, then spatial distribution, and participant distribution, should be reduced to a single type of distribution.

In favor of this argument, recall that, under the participant-distributive readings, the (sub)events must take place at different locations since the atomic (but also non-atomic) participants cannot simultaneously occupy the very same locations. This is exactly the reason why sentences with a singular participant, like (11) and (13), also do not yield the participant-distributive readings.

Intransitive Sentences

The following example illustrates *NumRed* in an intransitive sentence.

- (15) θə-ppð =əd **jiwəθ jiwəθ n təqfiʃθ**
 3F.SG-arrive.PFV = Dir one.F one.F Gen CS.girl
 Lit. Arrived **one one girl**.
“One girl/the girls arrived from different locations.”
“The girls arrived one by one/one at a time.”

(16) Event-Distributive Scenarios:

- a. *Spatial*: Dihya and Kenza left from your place together this morning. In the afternoon, Dihya arrived through your front door at the same time as Kenza, coming from a completely different place, for a different reason, separately, arrived through your back door. (Adapted from Matthewson 2000: 108)
- b. *Temporal*: Dihya arrived at 14 h00 and Kenza arrived at 14 h10.

The sentence in (15) is felicitous under the auspices of the temporal event-distributive scenario where the (sub)events involving one different girl are temporally separated, i.e., the girls arriving one at a time, as in (16b).

The intransitive sentence with the *NumRed* subject, in (15), naturally cannot yield the participant-distributive readings, since it has only one argument, marked for the distributive share by *Red*, *jiwəθ jiwəθ n təqfiʃθ* “one one girl”. There is no plural argument to be selected for the distributive-key, over which “one girl” could be

distributed (For comparison, see sentence (7) with *NumRed* in a subject position, in scenario (8a).).

However, (15) is felicitous under the spatial event-distributive scenario in (16a) where the (sub)events involving one different girl are simultaneous (have the same running times), but are spatially separated – the girls arrived at the same time but from different places. The intransitive sentence in (15), therefore, gives rise to both temporal and spatial event-distributive reading, contrary to the sentences with one singular argument in (11) and (13) which only give rise to the temporal event-distributive reading.

We have advanced the idea that the spatial separation must be made salient by the context, and that the participant separation is a relevant criterion for considering the spatial separation salient. Note that the spatial separation in (16a) is precisely the simultaneous participant separation (different (atom) girls arriving from different places). Importantly, however, note that in (15) the only participants in the event, girls, are marked as the distributive-share by *NumRed* (*jiwəθ jiwəθ n təqfiθ* “one one girl”). “One girl,” thus, must be distributed over the covert spatiotemporal argument – spatial locations and time units.

Recall that in Choe (1987), and subsequent analyses, the distributivity is a relation between the distributive-key and the distributive-share. For illustration, in (5), the entity marked as the distributive-share (*snaθ snaθ n təbwaʒin* “two two boxes”) is distributed over the set denoted by the other plural argument, the distributive-key (*θlaθa n təqfiθin* “three girls”). This is how the participant-distributive reading arises: “two boxes” are distributed to the atomic members of the set of three girls. While we descriptively use Choe’s terminology to talk about the NP marked with the *NumRed* as a distributive-share, and the unmarked NP as a distributive-key, we have contended that *NumRed* does not involve distributive dependency between the two arguments. Still using Choe’s terminology, we can say that the distributive-key NP serves as the criterion to separate the space units/locations over which the distributive-share is to be distributed. Importantly, in Choe’s approach, the distributive-key NP must be partitioned into the atomic members (so that each of the three girls carried two boxes in (5)). Contrary to this, our hypothesis is as follows. Spatial separation is always made via participant (atomic or non-atomic) separation.

In sentences where, descriptively, there is a suitable candidate for the distributive-key – plural argument not marked by *NumRed*, *θlaθa n təqfiθin* “three girls” in (5), and *snaθ n təbwaʒin* “two boxes” in (7), this argument *must* be used to separate the subevents spatially, by spatially dividing the plural participant into either atomic, or, importantly, non-atomic groups of participants. This predicts the non-atomic spatial distributive reading of (5), illustrated in (6b), where the non-atomic groups of girls participated in each subevent of carrying two boxes. This also predicts the atomic spatial distributive reading of (5), illustrated in (6a), where the atomic groups of girls participated in each subevent of carrying two boxes. Note that this reading is precisely the tantamount of the so-called participant-distributive reading.

Let us turn to the sentences that yield only temporal event-distributive reading, as (11) and (13), with one singular argument and one *NumRed* NP. These are sentences

without a suitable candidate for the distributive-key, since the only other non-marked argument as the distributive-share (by *NumRed*) is a singular DP *Dihya* in (11), and *θaxxamt* = *ənni* ‘the room’ in (13).

Since a singular participant cannot be at two different places at the same time, both participant and spatial-distributive readings are unavailable. The ‘participant criterion’ for permitting spatial separation seems to overrule the contexts in (12a) and (14a). This suggests that the presence of a singular NP seems to block the availability of a covert spatial argument to serve as a distributive-key. Therefore, the only allowed separation of subevents is temporal – the events involving *Dihya* carrying two boxes are separated/distributed over time intervals.

Finally, let us look at the sentences yielding, both spatial, and temporal, event-distributive readings, like the intransitive sentence in (15). As in (11) and (13), in (15) there is no suitable NP for the distributive-key either, since the only argument is the subject marked by *NumRed* as distributive-share, *jiwəθ jiwəθ n təqfiθ* ‘one one girl’. The intransitivity of the sentence naturally, thus, excludes the participant-distributive reading. The very important difference, however, between (11) and (13), yielding only temporal event-distributive reading, and (15), yielding, both spatial and temporal event-distributive readings, is that the former are transitive sentences (with two arguments), and the last is an intransitive sentence (with only one argument). That is, in (15), the question of a suitable NP for distributive-key is irrelevant, since it is a sentence with one argument only. The only NP being distributive-share (*NumRed* NP *jiwəθ jiwəθ n təqfiθ* ‘one one girl’), and the absence of a second NP, allow a covert spatial argument to be considered as a relevant, and moreover, a suitable candidate for the distributive-key. That is, the subevents involving ‘one girl arriving’ can be either simultaneous (spatial separation) or non-simultaneous (temporal separation).

To resume, the availability of the participant-distributive readings depends on the presence/absence of a plural NP in a sentence, which is not a distributive-share. If there is a singular NP in a sentence (not being the distributive-share), the integral argument must be considered as a potential distributive-key. Since it is not possible that the very same singular participant is in two events simultaneously, but at different places, spatial event-distributive reading does not arise. If there is not another NP at all (other than the distributive-share), as it is the case with intransitive sentences, then the spatial separation of events is possible in the same way as temporal separation: the events involving *n(umeral)* NP (numeral given by *NumRed* NP) can be separated either spatially, or temporally, i.e., distributed, either over spatial locations or time units. The nonavailability of the spatial event-distributive reading with transitive sentences with one singular argument, as well as the non-atomic participant-distributive readings of transitive sentences with two plural arguments, highlights that the spatial and participant distribution should merge to one and only kind of distribution. We argue that both cases should be considered as spatial distribution. The reason is the following: the participant-distributive scenario is the one where in each (sub)event there is an atomic participant. This is, however, not mandatory with *NumRed*, since, as we have seen in sentences (5) and (7), non-atomic partitioning of the group of participants is also acceptable (see

(6b) and (8b)). We, thus, conclude that the spatial distribution is obtained via participant distribution in the following way: if there is an NP which is not marked by *NumRed* (not distributive-share), then this NP must be plural, and it may be partitioned into atomic or non-atomic groups (see scenarios in (6b) and (8b)). If no such participant is available, that is, if the NP is singular, the spatial (and participant) event-distributive readings are unavailable. Consequently, only the temporal event-distributive reading arises. If, on the other hand, there is no other NP except the *NumRed* NP (distributive-share), the events involving *n* NP can be spatially, or temporally, separated. Consequently, both spatial and temporal event-distributive readings arise.

Atomicity and Exhaustivity

In the previous sections, we have discussed the non-atomic readings of sentences with *NumRed* in favor of reducing participant distribution to spatial distribution. The atomicity, as well as exhaustivity, are also adopted as a diagnostic for the universal quantification (Knežević 2015; Knežević and Demirdache 2018).

Here we explicitly illustrate the non-atomicity, and the non-exhaustivity of *NumRed* sentences. We compare both properties with the universal distributive quantifiers *mkul/each*.

In the examples of transitive sentences, with one simple numeral, and one *NumRed* in (5) and (7), we have seen, under the event-distributive scenarios in (6b) and (8b), that *NumRed*, unlike distributive universal quantifiers *mkul/each*, does not force distribution to atoms. That is, *mkul/each* force the distribution over atomic participants of the group of individuals denoted by an NP serving as a distributive-key. In other words, in the events of girls carrying two boxes, there must be one (atomic) girl in each (sub)event (see sentence (3), repeated in (18)). This is not the case with the sentence with *NumRed* in (4), repeated in (17).

- (17) rəfð-ənt θəqʃiʃin =ənni snaθ snaθ n təbwaʒin
 carry.PFV-3F.PL CS.girl.PL =DEM two.F two.F Gen CS.box.F.PL
 Lit. The girls carried **two two boxes**.
 “The girls carried two boxes **each**.”
 “The girls carried two boxes **at different times/places**.”

True in scenario (19)

True in scenario (20)

- (18) a. **mkul taqʃiʃθ** θə-rfəð snaθ n təbwaʒin
 each FS.girl 3.F.SG-carry.PFV two.F Gen CS.box.F.PL
 b. “**Each girl** carried two boxes.”

Not true in scenario (19)

Not true in scenario (20)

(19) Event-distributive scenario (non-atomicity):

Dihya, Anya, and Kenza need to help their neighbors bring up some boxes. Dihya and Anya carried two boxes together, while Kenza carried two boxes alone.

Unlike (18) with *mkul*, (17) with *NumRed* is felicitous in scenario (19), in which, in the (sub)events of carrying two boxes, there is a non-atomic group of girls (Dihya and Anya), and an atomic group (Kenza). This suggests that *NumRed* does not involve distribution over individuals, but over spatiotemporal units that may, but need not, involve atomic participants.

NumRed, unlike distributive universal quantifiers *mkul/each*, does not force exhaustive distribution either. That is, since *NumRed* does not distribute to participants but to spatiotemporal locations, it does not force exhaustivity of the group of (subject) participants. This is illustrated in scenario (20), in which not all the girls participated in the carrying event. This scenario can be described with the *NumRed* sentence in (17), but not with the *mkul/each* sentence in (18).

(20) Event-distributive scenario (non-exhaustivity):

The girls, Dihya, Anya, Kenza, and Tanina, must bring up some boxes. They carried two boxes each. O, of course, Tanina did not do anything but complain that she was too tired to carry.

The so-called non-atomicity and non-exhaustivity of distribution with *NumRed* is precisely the evidence that *NumRed* does not imply the universal quantification over individuals. The “non-atomic” and “non-exhaustive” event-distributive readings provide evidence that *NumRed* involves a plurality of events, temporally or spatially separated, with atomic or non-atomic participants in each subevent, and each involving the exact number of the entity provided by the numeral in *Num* NP.

Plurality Requirement

The issue with plurality requirement, addressed in the literature on distributive numerals (Balusu 2006; Knežević 2015; Bedar and Allawama 2024, among others), is whether, and to what extent, participants provided by *NumRed* NP must be different, or the same, in the events. That is, can the total of individuals denoted by NPs be exactly *n(umeral)* in the group of events (separated spatially or temporally), or the total of individuals denoted by NP in the group of events needs to be greater than *n*? In other words, does, for instance, *θata θlata n təqfifin* “three three girls” imply that the girls involved in the events must be three different girls per event, or three girls can always be the same, e.g., Dihya, Anya, and Kenza. Our Taqbaylit consultants only accept the contexts where in each subevent there are three different girls that participate. This means that the number of the participants in the

(whole) event is always greater than the number provided by the numeral in *NumRed* NP. Recall that distributive-shares must be indefinite (nonspecific) expressions of explicit quantity (Choe 1987). In Taqbaylit, it is not possible to use the (definite) demonstrative *-ənni* in a sentence with *NumRed*, as illustrated with the ungrammatical sentence in (21). By opposition, it is possible to use *ənni* in a sentence with a simple *Num*, illustrated in (22). This suggests that *Red* ensures the non-specific interpretation of the NP with which it combines, also permitting the pluralization of the participants denoted by the NP.

- (21) ***θlaθa θlaθa** n təqfiʃin = **ənni** rəfð-ənt
three Three Gen CS.girl.PL = DEM carry.PFV-3F.PL
snaθ n təbwaðin
two.F Gen FS.box.F.PL
Lit. ***The three three** girls carried two boxes.
- (22) **θlaθa** n təqfiʃin = **ənni** rəfð-ənt **snaθ** n təbwaðin
three Gen CS.girl.PL = DEM carry.PFV-3F.PL two.F Gen FS.box.F.PL
Lit. **The three girls** carried two boxes.

These examples suggest that different events (differentiated by *NumRed* in terms of different times/locations) also require different participants who are distributed. This descriptive fact supports the claim that *NumRed* is an event plurality marker, pluralizing, not only the verb but also its participants (denoted by the NP with which the *NumRed* combines). Note that, importantly, in sentences with *NumRed*, e.g., (5), the total number of participants given by *NumRed* NP (*snaθ snaθ n təbwaðin* “two two boxes”), will not depend on the number of the members of the other plural NP (*θlaθa n təqfiʃin* “three girls”), as is the case with sentences with distributive *mkul/each*, in (3).

In this section, we have shown that *NumRed*, involving distributivity, differs from the distributive universal quantifiers, such as *mkul/each* and shares some core properties with pluractional markers. First, all sentences with *NumRed*, unlike distributive universal quantifiers, yield temporal event-distributive readings, where the distribution is over time intervals. Some also yield spatial event-distributive readings: transitive sentences with two plural arguments and intransitive sentences. Participant distribution is a subcase of the spatial distribution. Second, participants in the described events do not have to be atomically partitioned. They can also be partitioned into groups or not partitioned at all as is the case with distributive universal quantifiers. This holds for all sentences with *NumRed* that have plural participants, not combining with the *NumRed*. Third, the set of entities, be it participants, locations, or times, over which the entity denoted by *NumRed* NP is distributed, does not need to be exhausted. This holds even in the case where the distribution is **over** atomic participants, which is not the case with the distributive universal quantifiers. Finally, the entity denoted by the *Num* NP combining with *Red* must be pluralized across the (sub)events as is the case in general with pluractional markers, pluralizing verbs and their participants. This implies that the total number of participants denoted by *n* in *Num* NP is always greater than *n*.

Analysis

In this section, we sketch an explicit proposal as to how the meaning of *Red* contributes to the meaning of sentences with numerals. In a nutshell, the claim is that sentences with *NumRed* NPs describe events where the number of participating individuals described by NP is not exactly n , but greater. The events described by the sentence must be constructed out of (sub)events involving exactly n individuals, but they must not be events of that kind themselves. We claim that (i) *Red* semantically combines with the numeral (n), and then with the NP, and (ii) there must be at least two events that involve an NP of the quantity n .

Theoretical Background

Theoretical accounts of the phenomena of distributive numerals or, more generally, distance distributivity (cf. Zimmermann 2002), can be understood as divided up into two theoretical camps. The first approach is based on the theories of universal quantification **over** individuals, and takes as an underlying assumption that distance-distributive quantifiers (distributive numerals) involve, each, a covert distributive operator (Choe 1987; Oh 2001; Zimmermann 2002). The second approach is based on the theories of event quantification or, more accurately, event plurality. The event-based approach splits into two directions: one approach still makes use of the universal quantifier whose restriction is always an event argument (Balusu 2006), and the other which implies semantics of the distributive element yields event plurality (Cable 2014; Knežević 2015).

The main shortcoming of the operator-based proposals is that they do not predict “intermediate” distributive readings of distributive numerals (those under which the atomic partition of a set is not required), and where the participants need not be exhaustively distributed over. Also, these analyses predict the exhaustivity and atomicity of a spatiotemporal argument. This is difficult to test since it is not clear, either theoretically, or empirically, what the exhaustive and atomic distribution over time units and spatial locations would mean.

The event-based analyses seem to avoid these problems. The main originality of Cable’s (2014) proposal, and the subsequent proposal in Knežević (2015), as opposed to previous analyses, is that they account for distributive readings in terms of sums of events which all involve an explicit number of participants (provided by a numeral). This avoids the problems encountered earlier, namely, the non-exhaustive distribution (over individuals), as well as the non-atomic participant-based distribution. That is, neither the number of participants in the event, nor their grouping into non-atomic or atomic groups, has an impact on the interpretation of a sentence. The semantics given to distributive numerals is large enough to cover all cases of event-distributive readings, including the so-called participant-distributive reading. Although they deal with the distributive numerals in Tlingit and Serbian, respectively, both authors additionally point out that the analyses can be

extended to pluractionals (in a more general way) in Kaqchikel and Serbian. This is in line with the hypothesis, defended in this chapter, that *NumRed* (distributive numerals) in Taqbaylit are pluractional markers.

Semantics of *Red*

Following Knežević's (2015) proposal for the distributive marker *po* in Serbian, we argue that *Red* in Taqbaylit is a marker of event plurality enforcing, rather weak truth conditions. This means that the situation described by the sentence containing a *NumRed* NP involves at least two events that are spatially or temporally separated, each of which must involve *n*NP(s) (where *n* stands for the numeral to which *Red* applies). The resulting sentence will be true in various scenarios, as long as the described events involving *n*NP(s) have different running times and/or running spaces. Essentially, *Red* contributes to the meaning of a sentence by conveying that there is a plurality of events, each involving *n*NP(s). We claim that (i) *Red*, semantically, combines first with a numeral (*n*) and then with an NP, and (ii) there must be at least two events that involve an NP of the quantity *n*. The semantics of *Red* is given in (23).

$$(23) \quad [[Red]] = \lambda n. \lambda Q_{\langle e, t \rangle}. \lambda P_{\langle e, et \rangle}. \lambda e. e \in *_{\varepsilon n} Q \ \& \ e \in \varepsilon n Q \ \&$$

To obtain the semantics below for *NumRed* NP, *Red* combines successively with the numeral (*n*) and the NP, in (24).

$$(24) \quad [[Red \ n \ NP]] = \lambda P_{\langle e, et \rangle}. \lambda e. e \in *_{\varepsilon n} NP \ \& \ e \in \varepsilon n NP \ \&$$

For illustration, let us see below how the analysis applies to the transitive sentences with one *NumRed* (in an object or in a subject position), and with double *NumRed* (in both subject and an object positions simultaneously).

The sentence with *NumRed* in an object position in (25a) will be assigned the semantics in (25b), where ε stands for a type of event involving *n*NP(s) (here, *two boxes*), and $*$ for a cumulative denotation of the predicate (Link 1983).

- (25) a.

θlaθa	n	təqʃiʃin	rəfð-ənt	snaθ	snaθ
three	Gen	CS.girl.PL	carry.PFV-3F.PL	two.F	two.F
n		təbwaðin			
Gen		CS.box.F.PL			

Lit. Three girls carried **two two boxes**.

- b. $\lambda e. e \in *_{\varepsilon 2boxes} \ \& \ e \in \varepsilon 2boxes \ \& \ *carry(e)$
 $\ \& \ box(x) \ \& \ Theme-sum(e)(x)$
 $\ \& \ girl(y) \ \& \ |At(y)| = 3 \ \& \ Agent-sum(e)(y)$
- c. There is an event constructed out of (sub)events each involving two boxes, and this is an event of three girls cumulatively carrying two boxes.

We can informally read (25b) as in (25c). On this proposal, the sentence in (25a) describes an event constructed out of (at least) two (sub)events involving three girls cumulatively (that is, together, in groups, or individually) carrying two boxes. Since here *Red* appears with the numeral “two” (the distributive-share), *Red* requires that each subevent involve exactly two boxes. Note in particular that the events described by (25a) are in $*\varepsilon_{2\text{boxes}}$, and they are sums of events in $\varepsilon_{2\text{boxes}}$, but they are not in $\varepsilon_{2\text{boxes}}$ themselves. They are, thus, constructed out of events in which exactly two boxes participate, but are not themselves like that. In other words, they must involve more than two boxes.

Let us suppose that the set of three girls is itself partitioned atomically and each girl atom individually carried two different boxes. Then, the so-called participant-distributive reading arises since the ensuing reading is on a par with distributing two boxes over individuals (girl) atoms. Supposing that the group of three girls is not partitioned into atoms and that the respective agent of each subevent is a group of three girls carrying simultaneously (together or in non-atomic groups of two and one) two different boxes (per subevent). The ensuing reading is equivalent to distributing events (of three girls carrying two boxes) over spatial locations, i.e., two (or more) carrying subevents are spatially, but not temporally, separated (that is, they are happening simultaneously), whereby the spatial event-distributive reading arises.

Finally, let’s suppose again that the group of three girls is not partitioned into atoms and that the agent of each subevent is a group of three girls carrying (together, in non-atomic groups, or individually) two different boxes per event, each of which has a different running time. This would be the case in a scenario where, for example, the three girls carry two (different) boxes (at least) twice during the course of the week. Thus, the temporal event-distributive reading arises.

Let us now see how the semantics in (23) contributes to the meaning of the sentence with *NumRed* in subject position, in (26).

- (26) a. **θlaθa θlaθa n təqʃiʃin rəfð-ənt snaθ**
 three three Gen CS.girl.PL Carry.PFV-3F.PL two.F
 n təbwaðin
 Gen FS.box.F.PL
 Lit. **Three three girls** carried two boxes.
 b. $\lambda e. e \in *\varepsilon_{3\text{girls}} \ \& \ e \in \varepsilon_{3\text{girls}} \ \& \ *\text{carry}(e)$
 $\& \ \text{sum}(e)(x) \ \& \ \text{Theme-sum}(e)(y)$
 c. There is an event constructed out of (sub)events each involving three girls, and this is an event of girls cumulatively carrying two boxes.

In (26a), since *Red* combines with the NP “three girls,” there must be at least two subevents, each involving a different set of three girls carrying two boxes.

First, let us suppose that the set of two boxes is itself partitioned atomically and each box atom was carried by three different girls. Then the so-called participant-distributive reading arises since the ensuing reading is tantamount to distributing three girls over individual (box) atoms.

Now suppose that the group (of two) boxes is not partitioned into atoms and that the respective patient of each subevent is a group of two boxes being carried together (per subevent), and simultaneously. The arising reading is spatial event-distributive where (at least) two carrying subevents are spatially, but not temporally, separated.

Finally, suppose that there are carrying events, such that in each subevent there are three girls carrying two boxes (two boxes per group of girls, or two boxes in total, in all subevents), doing so in different time intervals. Then the temporal event-distributive reading arises.

Finally, let us examine how the semantics in (23) applies to the sentence with *NumRed* in both subject and object positions in (27).

- (27) a. $\theta la\theta a$ $\theta la\theta a$ n $təqʃiʃin$ $rəf\delta\text{-}ənt$ $sna\theta$ $sna\theta$
 three three Gen CS.girl.PL carry.LPFV-3F.PL two.F two.F
 n $təbwa\theta in$
 Gen FS.box.F.PL
 Lit. **Three three girls carried two two boxes.**
 b. $\lambda e. e \in *e_{3girls} \ \& \ e \in e_{3girls} \ \& \ e \in e_{2boxes} \ \& \ e \in e_{2boxes} \ \& \ *carry(e)$
 $\& \ box(x) \ \& \ Theme\text{-}sum(e)(x) \ \& \ girl(y) \ \& \ Agent\text{-}sum(e)(y)$
 c. There is an event constructed out of (sub)events each involving three girls and two boxes, and this is an event of girls cumulatively carrying two boxes.

In (27a), *Red* combines with the NP “three girls” and with the NP “two boxes,” so there must be at least two subevents, each involving a different set of three girls, and a different set of two boxes. In other words, since “three girls” and “two boxes” are both marked for distributive-share (need to be distributed), they need to be participants in each subevent. That is to say, they must both be distributed to times intervals/locations. Consequently, the arising readings are event-distributive (spatial or temporal).

Consider the scenario in which there is a plural event constructed out of subevents, such that in each subevent the set of three girls and the set of two boxes participate, and the subevents are happening simultaneously, but are spatially separated. Then the spatial event-distributive reading arises.

Now, let us suppose that in each subevent there is a plural event constructed out of subevents, such that in each subevent the set of three girls and the set of two boxes participate at the same place, but at different time intervals. The arising reading is the temporal event-distributive.

The reader may apply the semantics of *Red* in (23) to calculate the meanings of transitive sentences with one singular argument, yielding only temporal event-distributive readings and with intransitive sentences, yielding spatial and temporal event-distributive readings (see Section “[Participant-Distribution Is Reducible to Spatial-Distribution](#)”). Note that, under the rubric of our analysis, *Red* contributes to the meaning of a sentence by conveying that there is a plurality of events, each involving *n*NP. We can say that the obtained plurality of events is constructed out of atomic events that must be separated, temporally, or spatially (be it with atomic or

non-atomic participant separation). However, the analysis does not account completely for this “separation” requirement. Notice that this separation requirement is a property of pluractionals (Lasersohn 1995; Matthewson 2000). When it comes to pluractionals, Lasersohn (1995) proposes that the (sub)events must not overlap in at least one of three criteria, as shown in (28).

- (28) Pluractional morphology in Lasersohn (1995)
- a. $V\text{-PA}(X) \Leftrightarrow \forall e, e' \in X [V(e) \& \neg \tau(e) \circ \tau(e')] \& \text{card}(X) \geq n$
 - b. $V\text{-PA}(X) \Leftrightarrow \forall e, e' \in X [V(e) \& \neg K(e) \circ K(e')] \& \text{card}(X) \geq n$
 - c. $V\text{-PA}(X) \Leftrightarrow \forall e, e' \in X [V(e) \& \neg \theta(e) \circ \theta(e')] \& \text{card}(X) \geq n$

(28) says that a pluractional verb (V-PA) holds of a group of events (X), if and only if it **holds of** each (sub)event e , which is a member of X. The remainder states that the (sub)events must have non-overlapping running times (28a), running spaces (28b), or thematic roles (28c). The non-overlapping/separation requirement is guaranteed by the meaning of a particular pluractional morpheme which is different in (28a), (28b), and (28c). As we have seen, we do not have different morphemes for different readings, namely, the morphemes involving separate running times (28a), separate locations (28b), or separate thematic roles (28c). It is *Red* that is responsible for any of those readings. In other words, the semantics of *Red* enforces rather weak truth conditions, which permit a variety of different event-distributive readings, spatially and/or temporally separated events (involving atomic or non-atomic participants). Therefore, the semantics of *Red* does not predict when the spatial (and participant) distributive readings are not available. This may be augmented in (23). Otherwise, we can speculate that this issue may be related to predicate types, (in) transitivity, argument structure, and (the alternative of) verbal reduplication, but a serious investigation of this broad phenomena, far beyond the topic of this chapter, is necessary to answer these questions.

We have shown that the semantics proposed for *Red* in (23) accounts for readings of sentences with *NumRed*. It correctly predicts that all sentences with *NumRed* yield event-distributive readings. More precisely, it claims that the truth of sentences with *NumRed*, in both types of scenarios (participant-distributive and event-distributive), derives from one and the same interpretation, involving a cumulation of events. In this way, we account for the meaning of sentences with *NumRed* without presuming an ambiguity.

Conclusion

We have shown that *NumRed* in Taqbaylit systematically involves event plurality. The sentences with *NumRed* yield distributive interpretations as a result of temporal and/or spatial separation of multiple events. Furthermore, we have shown how participant separation of events is reducible to the spatial separation, since different

spatial locations of the (groups of) participants are precisely the salient criterion for separation of relevant locations of the (sub)events: this rightly predicts that some sentences only yield temporal-distributive readings, since spatial separation of subevents is impossible due to the lack of a plural participant.

We have provided an analysis of *Red(uplication)* as a plurality marker. In a nutshell, we have claimed that the events described by the sentences with *NumRed* must be *constructed out of* events involving exactly n individuals, that is, must involve at least two (sub)events that are spatially or temporally separated, each of which must involve n NP(s). The resulting sentence will be true in various scenarios, as long as the described events involving n NP(s) have different running times and/or running spaces. Essentially, *Red* contributes to the meaning of a sentence by conveying that there is a plurality of events, each involving n NP(s).

The aim of our investigation of *NumRed* in Taqbaylit has been to contribute to the novel insights into distributive numerals and pluractionals across languages, as well to open the door to more investigations into reduplication, distributivity, and pluractionality in the Berber languages.

References

- Balusu R (2006) Distributive reduplication in Telugu. *Proc NELS* 36:39–53
- Bedar A, Allawama A (2024) Numeral reduplication as a pluractional marker in Jordanian Arabic. *Journal of Semitic Studies* fgad045. <https://doi.org/10.1093/jss/fgad045>
- Beghelli F, Stowell T (1997) Distributivity and negation: the syntax of each and every. In: Szabolcsi A (ed) *Ways of scope taking*. Springer Netherlands, Dordrecht, pp 71–107
- Bouzidi S (2012) La reduplication. Le cas de l'adverbe en berbère. In: Haßler G (ed) *Locutions et phrases: aspects de la prédication*. Nodus Publikationen, Münster, pp 175–185
- Cable S (2014) Distributive numerals and distance-distributivity in Tlingit (and beyond). *Language* 90(3):562–606
- Cabredo Hofherr P, Laca B (eds) (2012) *Verbal plurality and distributivity*. Walter de Gruyter, Berlin. [Linguistische Arbeiten 546]
- Cabredo Hofherr P, Pasquereau J, O'Meara C (2019) Event plurality in Seri. In: Johnson K, Alexander G (eds) *Proceedings of semantics of under-represented languages in the Americas 10*. CreateSpace Independent Publishing Platform, Scotts Valley, pp 1–16
- Champollion L (2012) Each vs. jewels: a cover-based view on distance distributivity. In: ALoni M, Kimmelman V, Roelofs F, Sassoon GW, Schulz K, Wetera M (eds) *Logic, language and meaning: 18th Amsterdam colloquium*, Amsterdam, The Netherlands, December 19–21, 2011, revised selected papers. Springer, Berlin/Heidelberg, pp 251–260
- Choe J-W (1987) *Anti-quantifiers and a theory of distributivity*. PhD thesis, University of Massachusetts, Amherst
- Cusic DD (1981) *Verbal plurality and aspect*. PhD thesis, Stanford University
- Gil D (1982) *Distributive numerals*. PhD thesis, University of California, Los Angeles
- Gil D (1988) Georgian reduplication and the domain of distributivity. *Linguistics* 26(6):1039–1066
- Gil D (1990) Markers of Distributivity in Japanese and Korean. *Japanese/Korean Linguistics* 1:385–393
- Gil D (1995) Universal quantifiers and distributivity. In: Bach E, Jelinek E, Kratzer A, Partee BH (eds) *Quantification in natural languages*. Springer Netherlands, Dordrecht, pp 321–362
- Kiss KE, Geröcs M, Zétényi T (2013) Preschoolers' interpretation of doubly quantified sentences. *Acta Linguistica Hungarica* 60(2):143–171

- Knežević N (2015) Numerals and distributivity in Serbian: at the syntax-semantics-acquisition interface. PhD thesis, Université de Nantes
- Knežević N, Demirdache H (2018) Universal quantification and distributive marking in Serbian. In: Kiss KÉ, Zétényi T (eds) *Linguistic and cognitive aspects of quantification*. Springer, New York/Singapore, pp 115–140
- Landman F (2000) *Events and plurality: the Jerusalem lectures*, vol 76. Kluwer Academic Publishers, Boston. <https://doi.org/10.1007/978-94-011-4359-2>
- Lasersohn P (1995) *Plurality, conjunction and events*. Kluwer, Dordrecht
- Link G (1983) The logical analysis of plurals and mass terms: a lattice-theoretical approach. In: Bauerle R, Schwarze C, von Stechow A (eds) *Meaning, use and the interpretations of language*. de Gruyter, Berlin, pp 303–323
- Matthewson L (2000) On Distributivity and Pluractionality. *Semant Linguis Theory* 10:98–114
- Newman P (1980) *The classification of Chadic within Afroasiatic*. Universitaire Pers, Leiden
- Newman P (1990) *Nominal and verbal plurality in Chadic*. Foris, Dordrecht
- Oh S-R (2001) Distributivity in an event semantics. *Semant Linguis Theory* 11:326–345
- Zimmermann M (2002) *Boys buying two sausages each: on the syntax and semantics of distance-distributivity*. PhD thesis, University of Amsterdam

Samir Ben Si Said

Contents

Introduction	597
Some Characteristics of the External Plural in Kabyle (Dialect of At Mangelat)	598
Uniform Plural Template in the Dialect of Dallet (At Mengelat Dialect)	601
CVCV Framework	601
Virtual Length and the Phonetic-Phonology Interface	602
The Initial Vowel of Nouns in Kabyle	604
Derivation of Kabyle Plural	605
Analysis of Diatopic Variation	606
Absence of Gemination in the Singular	609
Conclusion	611
References	611

Introduction

The range of studies on diatopic variation in Berber that are not simply descriptive – of the comparative type, with the aim of carrying out a diachronic reconstruction – but which seek to analyze the variation in a formal framework seems, at first sight, limited. Two important factors have led to the lack of interest shown in diatopic variation in Berber: on the one hand, the myth of a “Berber language,” shared by all Berber speakers, and on the other hand the extreme geographical, cultural, and linguistic fragmentation between Berber-speaking communities. However, the geographical fragmentation of these “islands” as well as the absence of any political organization (there has never been a Berber state in the modern era), and, consequently, of any sense of common cultural identity between the different communities, has led to great variation and linguistic distance between different Berber

S. Ben Si Said (✉)

Laboratoire Ligérien de Linguistique, Université de Tours, Tours, France

languages. Not only is there no mutual intelligibility between, say, Kabyle and Tuareg, but speakers of different varieties of Kabyle (depending on their geographical distance) do not understand each other either. This situation means that the few researchers working on Berber (compared to those working on other language families) focus their efforts, first, on the study of macro-variation: we compare Kabyle, Zenaga, Tuareg, Tachelhit, Siwi, etc. This granularity makes it difficult to zoom in on a particular language such as Kabyle to analyze the diatopic variation within it. It is with this need in mind that this work analyzes diatopic variation in the formation of nominal singulars, and plurals, in Kabyle.

This work attempts, to my knowledge for the first time, to describe the diatopic variation in Kabyle in a formal framework. It aims to address the question as to why diatopic variation is observable only in the formation of the nominal singular, whereas the plural is uniform in all dialects.

The first part of the paper deals with the formation of the plural. Focusing on Dallet's dialect (1982), the At Mengelat dialect, I propose that the size of the plural is uniform, namely, a sequence of 5 CV units and that this generalization applies to the whole of the Kabyle dialects. The second part of my analysis focuses on diatopic variation in the singular. It shows that the variation is due to the association of invariable lexical entries to different templates, that is, a given root selects a particular template, X, in dialect A, but chooses another template, Y, in dialect B. In this analysis, I adopt the idea that the root is vocalized, that is, that it contains vowels (Basset 1952; Bensoukas 2001, 2018; Cohen 1993; El Hamdi 2018, for a more recent and detailed analysis).

I argue that the observed dialectal variation is not due to differences in the computational system and does not concern the lexical entry that hosts the root (the segments and their linear sequence). Instead, I propose that what produces diatopic variation is the association of invariable roots with different templates from dialect to dialect. For example, an identical root for a given word, $\sqrt{\text{I}\beta\text{I}U}$ in all dialects, is lexically specified to be associated with the singular template A in the Azazga dialect, producing the surface form $\text{i}\beta\text{i}w$, but makes its singular in the template B in the Tazmalt dialect, producing $\text{i}\beta\text{i}$.

Some Characteristics of the External Plural in Kabyle (Dialect of At Mangelat)

There are three types of plurals in Kabyle Berber (The Kabyle data analyzed here comes from Dallet's (1982) Kabyle-French dictionary, which I transformed into a searchable database under Open Access. I specify that the variety of Kabyle in question is spoken in At Mangelat (Ain El Hammam), of which I am a native speaker.) (Chaker 1983; Mammeri 1986, among others) (Prasse (1974), with 11 types plural, offers a much more precise taxonomy – for Tuareg, but most of it carries over to Kabyle.): the external plural, characterized by the suffixation of -n to the singular form of the masculine and -in to the feminine/diminutive form. The second is called the internal or broken plural and is formed by an internal vocalic

alternation and a vowel –a- in the last vocalic position without any affixation. The third is known as the mixed plural, which is a combination of the two previous ones, that is, it is formed by the suffixing of -n and the vowel a- in the last vocalic position.

(1)

		Singular	Plural	Gloss
a.	Masculine	i-θβir	i-θβir-n	<i>Pigeon</i>
		a-xxam	i-xxam-n	<i>House</i>
	Feminine/diminutive	θ-i-θβir-θ	θ-i-θβir-in	<i>Pigeon (female)</i>
		θ-a-xxam-θ	θ-i-xxam-in	<i>Room</i>
b.		a-kanim	i-kanam	<i>Reed</i>
		a-jazið ^s	i-juzað ^s	<i>Chicken</i>
c.		a-zal	i-zilan	<i>Daylight</i>
		a-ð ^s að ^s	i-ð ^s uð ^s an	<i>Finger</i>

In addition to the suffixation of -n, the external plural can show on the surface what is traditionally interpreted as allomorphy of the suffix. This variation is illustrated under (2). The only stable property of the suffix seems to be the presence of a -n. (In the data presentation of this article, the initial vowel called “unstable” (diachronically prefixed) is shifted from the root, because it does not belong to the root, unlike the initial vowel called “stable” or “constant.” Only the construct state (CS) can inform us about the status of the initial vowel of nouns, whether it is radical or not: the stable/radical vowel holds and the unstable vowel falls, for example, alma (free state) ~ w-alma (construct state), a-rgaz (free state), and ~ w-rgaz (construct state) (for more details on the issue, see Basset (1945), Bendjaballah (2011), Ben Si Said (2020), and Idrissi (2000), among others).)

(2)

	Singular	Plural	Suffix	Gloss
a.	anz'að	anz'að-n	-n	<i>Hair</i>
b.	a-fus	i-fas-sən	-sən	<i>Hand</i>
c.	a-βəqqa	i-βəqqa-jən (It should be noted that the alternation of the initial vowel between singular and plural occurs in 97% of nouns (Basset 1945). Diachronically, it is a prefixed article which cannot, however, be dissociated from the root (Brugnatelli 1987, 1997; Stumme 1899; Vycichl 1957). This alternation plays no role in the demonstration below)	-jən	<i>Slap</i>
d.	a-jawa	i-jawa-wən	-wən	<i>Habitant of West Kabylia</i>
e.	alma	alma-θən	-θən	<i>Grassland</i>
f.	iɸər	afr-iwən	-iwən	<i>Wing</i>

(continued)

g.	izəm (In traditional grammar, schwa in Kabyle is inserted to avoid a sequence of three consonants (Chaker 1983: 43). In generative grammar, the behavior of schwa is analyzed as a vowel-zero alternation (Bendjaballah 2001; Ben Si Said 2014))	izm-awən	-awən	<i>Lion</i>
----	--	----------	-------	-------------

Before presenting my analysis and in order to consider that the only mark of the plural that can be determined is -n, it is then necessary to rule on the segments that are present in the plural but absent in the singular. They are identified under (3).

(3)

Singular	Plural	Element present in pl.	Suffix	Gloss
i-θβir	i-θβir-n	—	-n	<i>Pigeon</i>
a-βəqqa	i-βəqqa-j-n	j	-n	<i>Slap</i>
a-jawa	i-jawa-w-n	w	-n	<i>Habitant of West Kabylia</i>
alma	alma-θ-n	θ (Prasse (1974: 51) Prasse argues that the t, realized [θ] in Kabyle, that appears in the plural but is absent in the singular comes from a final *h lost in the singular, for example, sg. ebəki ~ pl. ibəkiten <sg. *ebakih)	-n	<i>Grassland</i>
izəm	izm-aw-n	aw	-n	<i>Lion</i>
ifər	afr-iw-n	iw	-n	<i>Wing</i>

As was pointed out in Chaker (1983: 90), there are two analytical options: either the segments that appear only in the plural are epenthetic, or they belong to the root but for some reason they are absent in the singular. I present here two arguments in favor of the latter option: the segments in question are part of the root.

The first argument is the unpredictability of elements that appear in the plural. Nouns that belong to the same singular template show different melodic elements in the plural. Here are some examples:

(4)

	Singular template	Singular	Plural	Epenthesis	Gloss
a.	CCCV	a-məçsa	i-məçsaw-n	w	Shepherd
		a-κərða	i-κərðaj-n	j	Mouse
		a-jəfki	i-jəfkiθ-n	θ	Milk
c.	VCC	isəm	ismaw-n	aw	Name
		izəm	izmaw-n	aw	Lion
		ixəf	axfiw-n	iw	End

The second argument is related to the forms of the feminine/diminutive singular. Indeed, elements that appear in the masculine plural but are absent in the masculine singular often appear in the feminine singular, as the following examples show.

(5)

Singular		Plural		Gloss
Masculine	Feminine/diminutive	Masculine	Feminine/diminutive	
a-məçsa	θ-a-məçsaw- θ	i-məçsaw-n	θ-i-məçsaw-in	Shepherd
a-ħəβj̥i	θ-a-ħəβj̥iw-θ	i-ħəβj̥iw-n	θ-i-ħəβj̥iw-in	Lean
i-mərzi	θ-i-mərziw-θ	i-mərziw-n	θ-i-mərziw-in	Break
a-βəqqa	θ-a-βəqqaj-θ	i-βəqqaj-n	θ-i-βəqqaj-in	Slap
i-qwi	θ-i-qwiθ-θ (It is not realized as geminate [θθ] but as a voiceless affricate [tʃ])	i-qwiθ-n	θ-i-qwiθ-in	Drum sticks
a-zzu	θ-a-zzuθ-θ	a-zzuθ-n	θ-a-zzuθ-in	Broom
a-hanu	θ-a-hanuθ (The word is attested in dialectal Arabic hanut and shows a t in the word final position)-θ	i-hanuθ-n	θ-i-hanuθ-in	Store

If we want to avoid multiple allomorphs in the plural, the unpredictability of segments present in the plural but absent in the singular allows only one analysis: these segments are part of the root.

Uniform Plural Template in the Dialect of Dallet (At Mengelat Dialect)

The generalization presented in this section will unify all plurals and show the floating melodic item in the singular. This generalization, whose core is a template of 5 CV units, is only valid if we adopt the hypothesis of the virtual length of peripheral vowels in Kabyle (Bendjaballah 2005). Before presenting the analysis of the plural, it is important, first, to introduce this hypothesis and the CVCV model.

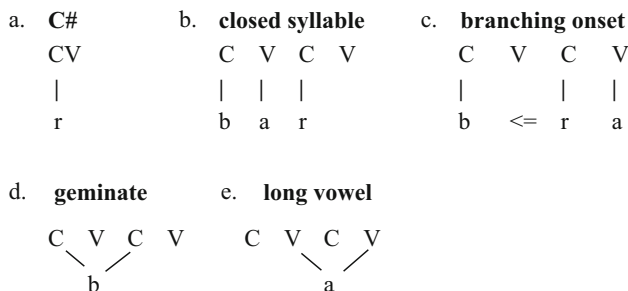
CVCV Framework

Autosegmental phonology (Leben 1973; Goldsmith 1976; McCarthy 1979) distinguishes between the skeletal level, which determines the number of segments in a sequence, and the syllabic level, which specifies their hierarchical relationships by means of a tree structure. The standard framework of Government Phonology (1990; Kaye 1990) is no exception. In this model, the syllabic level includes onset, rhyme, and nucleus constituents, which may be simple, or branching. Absent from the inventory are the syllable and the coda, which do not exist as such.

The CVCV approach (Lowenstamm 1996; Scheer 2004) stands out in that it attributes syllabic effects uniquely to lateral forces. In this framework, only two types of constituents are distinguished, non-branching onset and nuclei, which follow one another in strict alternation, and appear in the representations in the form of C and V for, respectively, onset and nuclei. Consonants are not specified in terms of onset or coda; it is the relations between them, or with vowels, that determine the syllabic structure of a linear chain.

In the CVCV model, a final consonant, a closed syllable, a branching onset, a geminate, and a long vowel will be represented as follows (I specify that a branching onset is always a TR group (T, obstructing/R, sounding) but the reverse is not true. There are TRs that are associated with a single syllable position, instantiating a contour structure, in the same way as the affricates (Lowenstamm 2003; Scheer 2004, 2019). Between the members of a branching (bipositional) onset, there is a lateral relationship in terms of their melodic content: the GI (infrasegmental government, Scheer 2004: § 14)):

(6)



Virtual Length and the Phonetic-Phonology Interface

In a modular conception of generative grammar (Fodor 1983), phonetics and phonology constitute two independent computational systems. The first works on articulatory and/or auditory objects. The second deals with abstract cognitive representations. (There are also approaches that consider phonetics and phonology to have a single computational system (cf. Flemming 2004).)

The purpose of the phonetic-phonological interface is to explain how the two modules communicate (Kingston 2007; Woohyeok 2012). How are cognitive units that exist only in the brain transformed into physical objects and vice versa?

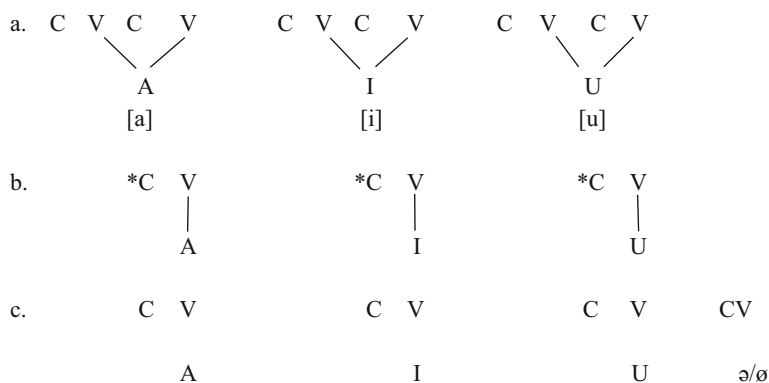
It has long been known that the relationship between the two modules is not always one-to-one: there can be distortion between phonological representations and their phonetic interpretations. Thus, a phonological object can have several phonetic interpretations and vice versa (Boersma 1998; Hamann 2011; Scheer 2014; among others). An example from Hamann (2011) is the vowel [e] in the Chukchi language (spoken in the Chukotka region of Siberia). This vowel has only one phonetic

realization but shows two different phonological behaviors: /e1/ lowers during vowel harmony, while /e2/ triggers it. Another example given by Scheer (2014) is the vowel *e* in Polish. This vowel shows two distinct behaviors: there are cases where *e* palatalizes (e.g., *lot* - *loci-e* “flight (plane) in the nominative case”), while in other cases, it does not produce this effect (*lot* - *lot-em* “flight (plane) in the instrumental case”). The work on Polish demonstrates that the two *e*’s are two different phonological objects but phonetically identical.

This distortion between phonetics and phonology has been developed in government phonology as *phonetic interpretation* (Harris and Lindsey 1995; Harris 1996). From this point of view, the phonetic interpretation of a phonological object is independent of the latter: it takes place after the end of the phonological computation and, in the same way that morpho-syntactic structure is transformed into phonological objects, the vocabulary of one (phonological) module is translated into that of another (phonetic) module by a spell-out operation.

In the CVCV framework, and according to the virtual length hypothesis (Lowenstamm 1991; for the Berber language: Bendjaballah 2005; Lahrouchi 2008; Lahrouchi and Ségéral 2010), the full vowels of the Kabyle have the representations under (7): the elements A, I, and U are associated with two vowel positions. When a full vowel is associated with a single vowel position, it is realized as schwa, or zero, depending on the phonotactic constraints required by the language (Bendjaballah 2005: 54).

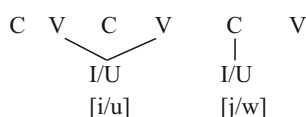
(7)



According to Bendjaballah (2005: 54), when a full vowel has access to only one vowel position, it remains floating (as in 7c) and cannot be associated with the skeleton level as in (7b). And she adds that if the phonotactic constraints of the language require that this position be realized, it is realized as a schwa; otherwise it remains uninterpreted phonetically.

I add that elements I and U are spelled out as *j* and *w*, respectively, when associated with a consonant position (Kaye and Lowenstamm 1984). Representations of I/U associated with the two types of syllabic constituents are given under (8).

(8)

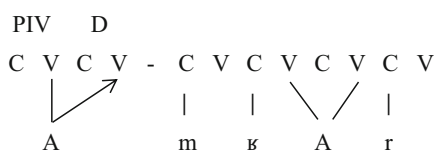


The Initial Vowel of Nouns in Kabyle

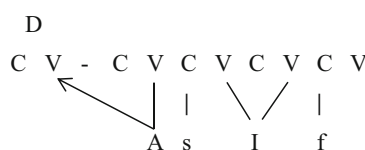
Almost all nouns in Kabyle (and in most Berber languages) start with a vowel. The latter can be radical or prefixal (Bendjaballah 2011; Ben Si Said 2020; Dell and Jebbour 1991; Lahrouchi 2011, 2013). In this article, to represent these two vowels, I adopt the proposals made in Bendjaballah (2011) and Ben Si Said (2020) for the same language and in the same theoretical framework. According to the authors, the initial radical vowel under (9b) is associated with the first V position of the root or template and the position V of determinant D. The lexical template for radicals with an initial vowel is “infirm”; it is during the derivation of the noun that this initial vowel of the radical is associated with a second V position and can be realized phonetically. As for the prefixed initial vowel (under 9a), it comes with its CV and spreads on the CV of D. (For more details on the motivation for the phonological representation of the initial vowel in Kabyle, see Bendjaballah (2011) and Ben Si Said (2020); and for the syntactic analysis of the initial vowel, see Achab (2003), Bendjaballah and Haiden (2013), Ennaji (2001), and Guerssel (1992), among others.) As discussed in Bendjaballah (2011) and as can be seen in (9), the radical of nouns with prefixal vowel is of the same size as the radical of nouns with radical initial vowel.

(9)

a. radical with a prefixed initial vowel

*amkar* ‘old man’

b. radical with a radical initial vowel

*asif* ‘river’

The following sections develop the hypothesis that the plural is reduced to a uniform template made of 5 CV, under the hypothesis of virtual length, and that the singular has different templates; melodic segments that appear in the plural but absent in the singular are floating because of the restriction of syllabic space offered by the template of the singular.

by singular templates, that is, the pronunciation of melodic objects can be constrained by the syllabic structure which does not provide enough space for them to be associated and, thus, to be realized at the surface level. This section shows that the diatopic variation of nouns in the singular works in the same way: a number of nouns show a variation in the singular across the different Kabyle dialects, whereas in the plural all dialects have the same size.

After having conducted several fieldwork studies in order to collect data from ten places, the major result is the invariability of the plural across the ten places of investigation: while the diatopic variation is general in the singular, the plural is uniform in all the varieties studied. Examples are given under (13).

(13)

Singular						Plural all varieties	Gloss
<i>Maatka</i>	<i>At Mengelat</i>	<i>Azazga</i>	<i>Yatafen</i>	<i>Akbou</i>	<i>Tazmalt</i>		
içən	içən	açniw	açniw	–	içən	açniw-n	<i>Twin</i>
anək	inək	–	inək	inɛi	–	inɛij-n	<i>Palate</i>
iβiw	iβiw	iβiw	iβiw	iβi	iβi	iβaw-n	<i>Bean</i>
–	a-ʃaʃiw	a-ʃaʃi	a-ʃaʃiw	a-ʃaʃi	a-ʃaʃi	i-ʃaʃiw-n	<i>Large hat</i>
a-jəfk	a-jəfki	a-jəfki	a-jəfk	a-jəfki	a-jəfki	i-jəfkiθ-n	<i>Milk</i>
–	iʃʃəw	iʃʃəw	iʃʃəw	iʃʃ	iʃʃi	aʃʃiw-n	<i>Horn</i>
i-nəʃ	i-nʃəw	a-nʃiw	i-nʃəw	a-nʃiw	i-nʃ	a-nʃiw-n	<i>Feather</i>

The invariability of the plural confirms the existence of a template, namely, that the plural by suffixation is not derived from the singular but formed independently by the association of root to the template of the plural. Indeed, if the external plural was derived by the suffixation of -n to the form of the singular, we would expect the same variation in the plural, as in the singular – as in column B of the table below (14). However, since the plural is invariant (column A), the concatenative hypothesis can be discredited.

(14)

Dialect	singular	plural		gloss
		A	B	
<i>Maatka</i>	i-nəʃ	} a-nʃiw-n	*i- nəʃ-n	feather
<i>At Mengelat</i>	i-nʃəw		*i- nʃəw-n	
<i>Azazga</i>	i-nʃiw			

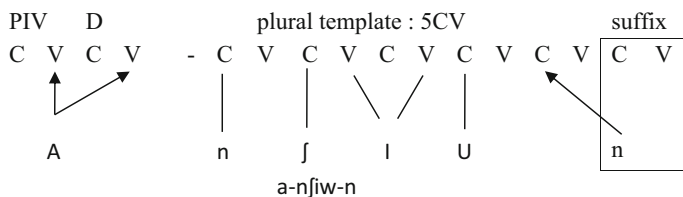
In addition, it can be observed from the table (13) that there are melodic items present in the plural that do not appear in the singular of some dialects and that these melodic items can be multiple for a given root. Thus, for each root, some dialects have missing segments in the singular. From the common plural a-nʃiwən “feather,” for example, we find in the Maatka dialect the sg. i-nəʃ (missing -iw), in the At Mengelat dialect the sg. i-nʃəw (missing the radical vowel i), and finally, the sg. i-nʃiw in Azazga (where all the melodic elements present in the plural are

realized). The form in Azazga is, therefore, the only one that is identical to the singular and the plural.

This variation of singular forms can be covered if, from a stock of singular templates that is common to all dialects, a given root selects its singular template idiosyncratically in each of the systems. The derivations, below, illustrate this perspective for the example of the plural of “feather” discussed.

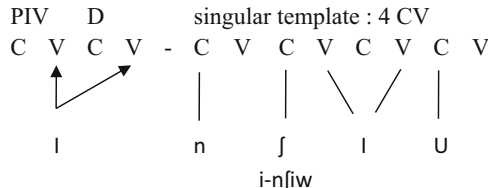
Let us consider, first, the plural [a-nʃiw-n] common to all dialects for the root $\sqrt{\text{nʃlw}}$. Its derivation is shown under (15). Note that in addition to the root the suffix -n is associated with the template of the radical to satisfy its size of 5 CV units.

(15)



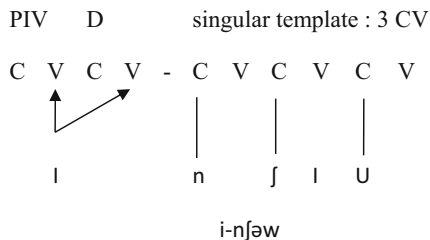
The singular form of this root observed at Azazga is produced, if associated with a singular template of 4 CV units.

(16)



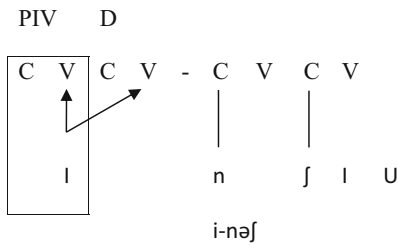
The singular form observed at At Mengelat dialect, on the other hand, is derived from a template of only three CV units.

(17)



In Maatka, finally, the singular is made on a template of two CV units only.

(18)



An important observation is that left-to-right association produces the right result in all three cases. Or, in other words, as the singular template does not provide enough space for all radical items to be present on the surface – that is, since it is reduced to three CV (At Mengelat), then to two CV (Maatka) – it is the rightmost melodic elements of the root that pay the price for this template’s reduction: first, in At Mengelat’s dialect, the vowel I can no longer be associated with two positions (so it is reduced to schwa), and then in Zoubga’s dialect, the U remains floating, in addition to the I associated with a single vocalic position, which is realized zero on the surface.

Absence of Gemination in the Singular

There are cases where the variation seems to indicate that some dialects do not tolerate gemination in the singular. Some such cases can be found under (19).

(19)

	Singular						Plural	Gloss
	Zoubga	At Mengelat	Azazga	Yatafen	Akbou	Freha		
a.	a-məʃʃaɖ	i-məʃʃaɖ-n	a-mʃaɖ	a-məʃʃaɖ	–	a-məʃʃaɖ		Thigh
b.		i-həʃʃir-n	i-hʃir	i-həʃʃir	i-hʃir	i-həʃʃir		Anger
c.	a-ʒəʃʃriɖ	i-ʒəʃʃriɖ-n	a-ʒriɖ	a-ʒəʃʃriɖ	–	–		Line
d.	a-fəssas	i-fəssas-n	a-fsas	a-fəssas	a-fsas	a-fəssas		Light
e.	i-jəllil	i-jəllil-n	i-jlil	–	i-jlil	i-jəllil		Poor

According to the table, it seems that in the dialects of Azazga and Akbou, systematically, there is no gemination of C2: all the other dialects have the geminate in C2 position. This variation, apparently based on the absence of gemination of a particular consonant, which is found in a particular dialect (Azazga and Akbou), is templatic: in all the dialects, the singular is made of five CV templates (as in the plural), this causes the median consonant to geminate to satisfy the template. On the other hand, in the dialect of Azazga and Akbou, the template of the singular is made of four CV, which means that it should not have a geminate consonant. To represent what has been said, let us consider the root $\sqrt{\text{hʃIr}}$ “anger” and first its plural, which as elsewhere is shared by all dialects. This plural is [i-həʃʃir-n], that is, it presents a

Conclusion

In this paper, I have developed an account of plural and singular formations in Kabyle. The analysis has argued that some lexical elements of roots do not surface in the singular because the template of their class does not leave enough space for them. It shows that the formation of the Kabyle plural is not as complex and irregular as it is presented in traditional grammars of the language. The template of the different plurals is uniform, that is, made of five CV units. Contrary to the plural, there is no restriction regarding size in the singular. This is why the formation of the plural seems to show distortions with respect to the singular:

The nominal size : singular, *different sizes*, vs. plural, *one size*

This asymmetry means that melodic elements may appear in the plural but are absent in the singular. Based on original data collected from different varieties of Kabyle, I have also demonstrated that the generalization proposed can be extended to all Kabyle dialects. The diatopic variation in Kabyle can, thus, be reduced to the following locus: the observable variation in the singular is due to the association of a root to different templates in different dialects: dialect A associates a root X to a template Y, and dialect B associates the same root (X) to a template Z.

References

- Achab K (2003) Alternation of state in Berber. In: Lecarme J (ed) *Research in Afroasiatic grammar* II. John Benjamins Publishing, Amsterdam, pp 1–19
- Basset A (1945) Sur la voyelle initiale en berbère. *Rev Afr* 86:82–88
- Basset A (1952) *La Langue Berbère*. Oxford University Press, London
- Ben Si Said S (2014) De la nature de la variation diatopique en Kabyle: étude de la formation des singuliers et pluriels nominaux. PhD dissertation, University Côte d’Azur, Nice
- Ben Si Said S (2020) La voyelle initiale des noms et l’état d’annexion en kabyle (berbère). *Rev Can Ling* 65(2):155–180
- Bendjaballah S (2001) The negative preterit in Kabyle Berber. *Folia Ling* 34:185–223
- Bendjaballah S (2005) Longueur phonologique des voyelles en kabyle. *Etudes Doc Berbères* 22: 47–69
- Bendjaballah S (2011) Note sur la voyelle initiale stable en kabyle. In: Mettouchi A (ed) *Parcours berbères, Mélanges offerts à Paulette Galand-Pernet et Lionel Galand pour leur 90e anniversaire*. Rüdiger Köppe Verlag, Cologne, pp 417–434
- Bendjaballah S, Haiden M (2008) A typology of emptiness in templates. In: Hartmann J, Hegedús V, Riemsdij HV (eds) *Sounds of silence: empty elements in syntax and phonology*. Elsevier, London, pp 21–57
- Bendjaballah S, Haiden M (2013) The representational anomalies of floating markers: light prepositions in Taqbaylit of Chemini. In: Biberauer T, Roberts I (eds) *Challenges to linearization*. Mouton, Berlin, pp 331–376
- Bensoukas K (2001) Stem forms in the nontemplatic morphology of Berber. PhD dissertation, Mohammed V University, Rabat
- Bensoukas K (2018) Against the consonantal root. *Asinag* 13:53–90
- Boersma P (1998) *Functional phonology. Formalizing the interactions between articulatory and perceptual drives*. Holland Academic Graphics, The Hague

- Brugnatelli V (1987) Deux notes sur l'état d'annexion en berbère. In: Jungrathmayr H, Mueller WW (eds) *Proceeding of 4th international Hamito-Semitic congress*. John Benjamins Publishing, Amsterdam, pp 349–359
- Brugnatelli V (1997) L'état d'annexion en diachronie. In: Mauro T, Bausi A (eds) *Afroasiatica Neapolitana. Contributi presentati all'8° Incontro di Linguistica Afroasiatica*. Istituto Universitario Orientale, Napoli, pp 139–150
- Chaker S (1983) *Un parler berbère d'Algérie (syntaxe)*. Publications de l'Université de Provence, Provence
- Cohen D (1993) Racines. In: Drouin J, Roth A (eds) *A la croisée des études libyco-berbères. Mélanges offerts à Paulette Galand-Pernet and Lionel Galand*. Geuthner, Paris, pp 61–175
- Dallet J-M (1982) *Dictionnaire Kabyle-Français*. SELAF, Paris
- Dell F, Jebbour A (1991) Phonotactique des noms à voyelle initiale en berbère (chleuh de Tiznit, Maroc). *Ling Anal* 21:119–147
- El Hamdi F (2018) On Tachelhit root structure and its implication for the organization of the lexicon. PhD dissertation, Mohammed V University, Rabat
- Ennaji M (2001) The construct state in Berber. *Stud Ling Sci* 31(2):55–72
- Flemming E (2004) Contrast and perceptual distinctiveness. In: Hayes B, Kirchner RM, Steriade D (eds) *Phonetically based phonology*. Cambridge University Press, London, pp 232–276
- Fodor JA (1983) *The modularity of mind: an essay on faculty psychology*. MIT Press, Cambridge
- Goldsmith J (1976) *Autosegmental phonology*. PhD dissertation, MIT
- Guerssel M (1992) On the case system of Berber. *Can J Ling/Rev Can Ling* 37(2):175–196
- Hamann S (2011) The phonetics-phonology interface. In: Kula NC, Botma ED, Nasukawa K (eds) *Continuum companion to phonology*. Continuum International Publishing, London/New York, pp 202–224
- Harris J (1996) Phonological representations are redundancy-free and fully interpretable. In: Laks B, Durand J (eds) *Current trends in phonology models and methods*. European Studies Research Institute, Manchester, pp 305–332
- Harris J, Lindsey G (1995) The elements of phonological representation. In: Durand J, Katamba F (eds) *Frontiers of phonology atoms, structures, derivations*. Longman, London, pp 34–79
- Idrissi A (2000) *Towards a root-and-template approach to shape-invariant morphology*. PhD dissertation, UQAM
- Kaye J (1990) Government in phonology: the case of Moroccan Arabic. *Ling Rev* 6:131–159
- Kaye J, Lowenstamm J (1984) De la syllabicité. In: Dell F, Vergnaud J-R (eds) *Forme sonore du langage: structure des représentations en phonologie*. Hermann, Paris, pp 123–159
- Kaye J, Lowenstamm J, Vergnaud J-R (1990) Constituent structure and government in phonology. *Phonology* 7(2):193–231
- Kingston J (2007) The phonetics-phonology interface. In: Lacy PD (ed) *The Cambridge handbook of phonology*. Cambridge University Press, London, pp 401–434
- Lahrouchi M (2008) A templatic approach to gemination in the imperfective stem of Tashlhiyt Berber. *Stud Afr Ling* 37(1):21–60
- Lahrouchi M (2011) Têtes et gabarits. *Etudes Doc Berbères* 29/30:255–274
- Lahrouchi M (2013) Templates, markers and syntactic structure in Tashlhiyt Berber. *Lingua* 133:53–72
- Lahrouchi M, Ségéral P (2010) Peripheral vowels in Tashlhiyt Berber are phonologically long: evidence from Tagnawt. *Brill's Annu Afroasiatic Lang Ling* 2:202–212
- Leben W (1973) *Suprasegmental phonology*. PhD Dissertation, MIT
- Lowenstamm J (1991) Vocalic length and syllable structure in Semitic. In: Kaye AS (ed) *Semitic studies: in honor of Wolf Leslau on the occasion of his 85th birthday*. Harrassowitz Verlag, Wiesbaden, pp 949–965
- Lowenstamm J (1996) CV as the only syllable type. In: Laks B, Durand J (eds) *Current trends in phonology models and methods*. University of Salford, Salford, pp 419–441
- Lowenstamm J (2003) A propos des gabarits. *Rech Ling Vincennes* 22:7–30

- McCarthy John J (1979) Formal problems in Semitic phonology and morphology. Phd dissertation, MIT, Cambridge
- Mammeri M (1986) Précis de grammaire berbère (kabyle). Awal, Paris
- Prasse K-G (1974) Manuel de Grammaire Touaregue (Vol. VI–V: Nom). Akademisk Forlag, Copenhagen
- Scheer T (2004) A lateral theory of phonology: what is a CVCV, and why should it be? Mouton, Leiden
- Scheer T (2014) Spell-out, post-phonological. In: Szpyra-Kozłowska J, Cyran E (eds) Crossing phonetics-phonology lines. Cambridge Scholars, Newcastle upon Tyne, pp 255–275
- Scheer T (2019) Sonority is different. *Stud Polish Ling* 14:127–151
- Stumme H (1899) Handbook of the Schilhschen by Tazerwalt. Hinrichs, Leipzig
- Vycichl W (1957) L'article défini du berbère. In: Memorial Andre Basset. A. Maisonneuve, Paris, pp 139–146
- Woohyeok C (2012) On the relation between phonetics and phonology. *Linguistic Res* 29(1): 127–156

The Functions of Independent Personal Pronouns in Berber

29

Valentina Schiattarella

Contents

Introduction	616
Typology of Independent Personal Pronouns: Forms and Functions	617
Independent Personal Pronouns in Berber	618
Forms of Independent Personal Pronouns	618
Functions of Independent Personal Pronouns in Berber	622
Independent Personal Pronouns in Siwi	623
Morphology and Syntax of Independent Pronouns	623
Information Structure Functions	627
Conclusions	632
References	633

Abstract

The paper investigates independent personal pronouns in Berber, outlining their typological profile and their syntactic and pragmatic functions.

These pronouns are a category that, with very few exceptions, is present in the grammatical system of all the languages of the world. There is a great variety of functions and forms linked to them, according to the source they originated from. All Berber languages have independent pronouns, but not all of them share the same grammatical features. Moreover, their syntactic and pragmatic functions are rarely found in the available grammars, which makes comparison difficult. Following an initial overview, the paper focuses on the functions of these pronouns in Siwi, a Berber language spoken in Siwa (Egypt).

Abbreviations

AOR	aorist
COMP	complementizer

V. Schiattarella (✉)
University of Hamburg, Hamburg, Germany
e-mail: vale.schiattarella@gmail.com

DO	direct object
F	feminine
IDP	independent pronoun
IMP	imperative
IO	indirect object
IPFV	imperfective
IRR	irrealis
M	masculine
PFV	perfective
PL	plural
POSS	possessive
PRAGM	pragmatic relevance marker
REL	relative marker
SG	singular
VN	verbal noun

Introduction

The use of non-bound or independent personal pronouns in a language with obligatory marking of subject affixes (person indexes) on the verb carries specific syntactic and pragmatic functions: Berber languages (Afroasiatic phylum) fall under this group. The same goes for lexical nouns as there are, indeed, long sections of discourse where neither independent pronouns nor lexical nouns appear co-referent to subject affixes (see Galand (1964), Mettouchi (2005), Mettouchi (2018), or Mettouchi and Schiattarella (2018), for a comparison between Kabyle and Siwi).

This paper addresses the forms and functions of independent personal pronouns only, and it is mainly based on data from Siwi, a Berber language spoken in Egypt, though a broad overview of these pronouns in other languages in general, and in several Berber languages in particular, is provided in the first part of the article.

Berber languages are spoken in Africa from the Mediterranean to the Sahel and from the Atlantic Ocean to Siwa, its easternmost point. Siwi is the only Berber (Amazigh) language spoken in Egypt by more than 30,000 people. Its speakers are nearly all bilingual (with Arabic, which can be Egyptian, mainly Cairene, or Bedouin).

The second part of the paper is based on data collected by the author since 2011 in the oasis of Siwa, but it also makes use of scattered elicitation sessions. All examples provided here come from this corpus, unless mentioned otherwise. The corpus is varied and considers different kinds of topics and genres: mainly narrations but also descriptions and dialogues. Vycichl (2005), Naumann (2012), Souag (2013), and Schiattarella (2017) are the most recent volumes on the Siwi language. Papers and older sources will also be cited, when relevant.

The paper is organized as follows: after a presentation on the typological aspects of independent pronouns in Section “[Typology of Independent Personal Pronouns](#):

Forms and Functions”, the forms attested in other Berber languages are analyzed in Section “**Independent Personal Pronouns in Berber**”, alongside their specific morpho-syntactic features, when available. In Section “**Independent Personal Pronouns in Siwi**”, the system of Siwi independent pronouns is presented, as well as their syntactic and pragmatic functions.

Typology of Independent Personal Pronouns: Forms and Functions

Independent personal pronouns are a closed-set category present in the vast majority of languages. Several scholars have discussed the issues of the very definition and delimitation of these pronouns. Pronouns do not always replace nouns or noun phrases, but the term is widespread and it is often used in its “broad sense” (Haspelmath 1997: 10). Scholars have also often pointed out the distinction (since Benveniste (1947)) – within the group of personal pronouns – between first- and second-person pronoun on the one hand and third-person pronouns on the other, the two sets having completely different functions. Studies have also focused on the genesis of these pronouns; their source of grammaticalization; their politeness or honorific functions; their gender distinctions (quite frequent in languages spoken in Africa like those languages belonging to Afro-asiatic, Niger-Congo, and Khoisan); the difference of pronominal number reference structure, if compared to nominals (we \neq I + I); and the distinction between inclusive and exclusive first-person plural pronouns (Bhat 2004, 2013; Cysouw 2013; Daniel 2013; Helmbrecht 2002, 2013; Heine and Song 2010, 2011; Siewierska 2013, among others).

The main point about the separation of first- and second-person pronouns from third-person pronouns is the fact that the former are not strictly referential but are only meant to indicate speech roles (speaker and addressee, also called “speech-act participants” (Haspelmath proposes the term “locuphoric” for one/two person pronouns, instead of “speech-act participants” (dlc.hypotheses.org/2430))), a relationship that is not stable and changes every time the speaker becomes the listener, such as in a conversation. Languages with no independent personal pronouns are rare but attested (Bhat 2004: 26, cites Mbay (Nilo-saharan)), as are languages with no personal pronouns at all, even though the nominals used instead of the pronouns (to express social status, politeness) could be regarded as pronouns. Personal pronouns are also often discussed with regard to the possibility for first-person plural pronouns to provide an exclusive or inclusive interpretation, which is used by Bhat (2004) as a way to explain why plurality in nominals is semantically distinguished from plurality of pronouns, which is also better understood as a conjunction of speech roles (Daniel 2013). Most languages do not make any distinction, but others (rather uncommon in Africa and Eurasia) have different forms for inclusive and exclusive pronouns. A small number of languages have a separate form for the inclusive, while the exclusive has the same form of the first-person singular (Cysouw 2013).

Gender distinction for third-person only (singular and/or plural) is the most widespread cross-linguistically. Additionally, other languages also make gender distinction for second and, in rarer cases, first person (Siewierska 2013). Gender distinction in second-person pronouns are rare and mainly belong to the Afroasiatic phylum. Those that mark it for first-person are even rarer and usually mark it only in the plural (Bhat 2004: 109–110).

To account for the distinction made between first- and second-person pronouns and third-person pronouns, Bhat seems to establish a typology of languages where those that have third-person pronouns as their demonstratives are more likely to show gender agreement (Bhat 2004: 134–136). Another aspect of personal pronouns studied in the literature is their genesis and the grammaticalization path. In Heine and Song (2010, 2011), only independent personal pronouns are considered. The authors identify five main sources for these pronouns: (1) nominal concepts, (2) spatial deixis, (3) identifiers, (4) pluralization, and (5) shift in deixis (2010: 121; 2011).

Personal pronouns are not exclusively studied in linguistics. Their use has been investigated in philosophy and literature, as well as in language acquisition and psycholinguistics, or in virtue of their relationship with gesture, and other deictic devices.

As for their function, independent personal pronouns are often defined by their “supportive role. Their function is merely to emphasize (or contrast) the involvement of speech roles in the events or states that the predicate denotes”, and “...independent pronouns are used only when special meanings like disjunction, emphasis or contrast are to be indicated” (Bhat 2004: 24–25).

Nevertheless, what at first glance appears to be a universal category can only be understood in a language-based description and, more importantly, together with other referential means, made available by the language. As stated in Mithun (1990: 361),

The function of pronouns might be assumed to be universal among languages. The use and interpretation of particular forms within a language, however, depend significantly on the full repertoire of referential devices with which they contrast. In English, third-person pronouns alternate with full lexical noun phrases and reflexives. In many other languages, free pronouns appear only for contrast or emphasis, while bound pronouns appear obligatory in every clause, whether additional nominals are present or not. Many languages lack regular third-person pronouns altogether; reference is indicated by means of full noun phrases, by demonstratives, or not at all.

Independent Personal Pronouns in Berber

Forms of Independent Personal Pronouns

The main distinction made by Galand (2010: 96) in Berber is between determination carrier pronouns, called in French by the author *pronoms support de détermination*, and personal pronouns.

Personal pronouns are then divided into (1) independent and (2) affixes or clitic pronouns. Moreover, a further distinction is made between (1) independent personal pronouns (In various other grammars of Berber varieties, independent personal pronouns are also called independent pronouns (in French *pronoms indépendants, isolés ou autonomes*); subject pronouns (despite the fact that their function is not that of subject, as already remarked by Basset in 1952 (29–30)) are called emphasis pronouns (*pronoms d'insistance*, Penchoen 1973: 73).), which are grouped together with direct pronouns (*pronoms régime direct*) (These pronouns are mainly, but not exclusively, direct object pronouns. Many Berber languages have two series of these pronouns (Brugnatelli 1993, 1998; Kossmann 1997b).), and (2) indirect pronouns (*pronoms régime indirect*), used for the indirect object or with prepositions, kinship terms, and nouns (together with the preposition *n* “of”) (1966, 2010). Independent pronouns are placed in the first group because of the formal similarities they share with the *régime direct* series (They also correspond to nouns in free state, while the rest of the pronouns correspond to nouns in the annexed state (Galand 1966: 289–298).), except for the first-person singular. As far as their morphology is concerned, Galand points out: “Les pronoms autonomes sont dérivés des affixes par adjonction d’éléments divers [...], ce qui met en évidence l’emploi souvent affectif ou expressif des pronoms autonomes. Aux autonomes reviennent les rôles de prédicat d’une proposition nominale (avec ou sans rhématisation) [...], d’indicateur de thème [...]; ils ont aussi une fonction d’appel (“vocative”) à la deuxième personne” ([Independent pronouns derive from affixes, with the addition of different elements [...], underlining the affective or expressive use that independent pronouns often have. Independent pronouns are predicates of a nominal proposition (with or without focalization) [...] topic [...]; second person pronouns can also be used for vocative.] All translations from the French, throughout the paper, are mine.) (Galand 2010: 111). See also Prasse (1972: 180–184) who comes to similar conclusions and offers a detailed morphological analysis of Tuareg (Ahaggar) pronouns. Chaker (2004: 48) also reconstructs independent personal pronouns and recognizes the similarity with personal affixes (*régime direct*), preceded by “the support element” *n* (except for second person) and followed by gender and number morphemes (*t* and *m* for feminine, *n* for plural).

Personal subject affixes have a particular status because they are obligatorily marked on the verbs, even in the case of the presence of a co-referent noun; and they also follow a different morphology compared to other affixes (Galand 2010: 104–109). That is why they are not listed among the other personal pronouns above.

Table 28.1 is not meant to be exhaustive but tries to give an overview of the forms of these pronouns in some Berber languages. In order to be as representative as possible, one language for each block used by Kossmann (2020: 284–285) to classify Berber languages, was chosen. This already rich picture is further complicated by the quite important internal variation of the main big varieties (i.e., Senhaja, Renisio (1932), and Gutova (2021); Lafkioui (2007, 2021); Kossmann (2011); Benrejda (2015); Tasahlit, Garaoun, personal communication, just to give a few examples). The paradigm developed by Paradisi for Awjila is also slightly different from the one proposed by Beguinot (1921), as mentioned by van Putten (2014: 107).

Table 28.1 Independent personal pronouns^a

	Zenaga	Tahaggart	Tashelhyt	Ghomara	Kabyle	Zenati Shawiya	Zenati Zwara	Ghadames	Awjila
<i>Singular</i>									
1	ni?K	nāk, nākkunan	nkki, nkkin	nekk, nekki, nekkīn, nekkineṭ	nekk, nekki, nekkini	n.čč	nəf	nāšš, nāššan	nək, nək
2M	kaK	kāy, kāyyunan	kiyyi, kiyyen	kežž, kegi, kegin, keḡineṭ	kečč, kečči, keččini	š.kk	ʃakk	šəgg, šəgəgən	ku, kù
2F	kaM ^m	kām, kāmūnan	km̄mi, km̄min	kem̄m, kem̄mi, kem̄min, kem̄mineṭ	kem̄m, kem̄mi, kem̄mini	š.m	ʃamm	šām̄m, šām̄man	kom
3M	nəttā, əntā	ənta	nfta, nftan	netta	netṭa	n.fta	nətta	nitto, itto	nəttin, əntin
3F	nəttāʔhād, ntāʔhād		nftat	nettata	netṭat	n.ftat	nəttat	nittāt, ittāt	ənnāt
<i>Plural</i>									
1M	nəkni	nākkāneṭ	nk ^w .ni, nkwnin	nukna	nekkni	n.šin	nəʃnin	nākkānen	nəkkāni
1F	nəKnaʔN ^ʔ ād	nākkānāteṭ	nk ^w .nti, nkwn̄tin		nekknti				
2M	nətni	kāwāneṭ	k ^w .ni, kwn̄in	kunna	kunwi	k.niun	nək̄nim	šək̄wēn	kəmmim
2F	nətnaʔnm ^ʔ ād	kāmāteṭ	k ^w .nin̄ti, kwn̄n̄tin		kunnem̄ti	k.niun̄ti ?	nək̄nimat	šək̄matēn	kəmmimat
3M	nəhni	əntāneṭ	n̄tni, n̄tinin	n̄hma	n̄utni	n̄ihin	nətnin	əntənēn	nəhin, ənhin
3F	nəhnaʔN ^ʔ (ād)	əntānāteṭ	n̄tn̄ti, n̄tn̄tin		n̄utenti	n̄ih.n̄tin	nətn̄inat	ənt̄natēn, ənnatēn	nəhin̄at, ənhin̄at

^aZenati is a language continuum that spreads from Morocco to Libya. Due to its vastness, the table contains two varieties (Zwara in Libya and Shawiya in Algeria). The transcription used in this table is the one used by the authors cited. For more information about phonological and morphological variation, the reader is referred to the original works (see footnote 7). k > č, š (first and second person) is a diagnostic feature of Zenati. For a more in-depth study, see Kossmann (1999: 173–182)

Some varieties have two forms for each person. The longer one is usually considered to be, at least originally, more emphatic than the shorter. The augmentative suffixes are the same found with different deictics (see Kabyle examples in Chaker (2004: 51))

From a typological point of view, the kind of display of gender distinction that Berber makes is quite rare: if no language has gender distinction in the first-person singular, many mark it in the first-person plural, and all of them mark it for the second-person singular. Those that mark gender distinction in the first-person plural also mark it in all other persons and numbers, in the plural.

Nevertheless, the majority of languages mark gender distinction only for the second and third person, singular and plural. Interestingly, Tuareg varieties (except Ghat in Libya) mark first-person plural gender agreement, but they lack gender distinction for the third-person singular. El Fogaha (Libya) only marks gender in the singular- and third-person plural. Finally, a few languages do not mark gender distinction in the plural, such as Ghomara (Morocco) and Siwa, as we will see in Section “[Morphology and Syntax of Independent Pronouns](#)”. Table 28.2 lists the main Berber languages. As mentioned before, the picture is even more varied if one looks at dialectal variation: for example, not all Rif varieties mark gender distinction in the first-person plural (Kossmann 2000; Lafkioui 2007: 116) and not all Senhaja varieties mark gender distinction in the second- and third-person plural (Gutova 2021).

Table 28.2 Gender marking in independent pronouns

Berber varieties ^a	1SG	1PL	2SG	2PL	3SG	3PL
Zenaga, Tetserret, Tamazight, Tashelhyt (Agadir), Rif (Eastern), Ayt Seghroushen, Kabyle (Central), Ouargla, Douiret, Ghat (Tuareg, Libya), and Nafusi	—	+	+	+	+	+
Senhaja, Rif (Western and central), Figuig, Chenoua, Beni Snous, Mzab, Shawiya, Tarighit (Temaçine), Timimum, Touat, Harakta, Chennini, Tamazret, Djerba, Sened, Ghadames, Zwara, Sokna, and Awjila	—	—	+	+	+	+
Tuareg: Tayert (Niger), Tahaggart (Algeria), Tamaghit and Tudalt (Burkina Faso), and Tamashek (Mali)	—	+	+	+	—	+
El Fogaha	—	—	+	—	+	+
Ghomara, Siwa	—	—	+	—	+	—

^aData for Tables 28.1 and 28.2 belong to the following sources: Zenaga (Taine Cheikh 2005), Tetserret (Lux 2013), Tamazight (Taifi 1991), Tashelhyt (Anab 2000), Eastern Riffian (Kossmann 2000; Lafkioui 2007), Ayt Seghroushen (Kossmann 2017), Kabyle (Mettouchi 2005), Ouargla (Basset 1893), Douiret (Gabsi 2003), Ghat (Nehilil 1909), Nafusi (Beguinet 1942), Senhaja (Renisio 1932; Gutova 2021), Western and Central Riffian (Lafkioui 2007), Figuig (Kossmann 1997a), Chenoua (Laoust 1912), Beni Snous (Destaing 1907), Mzab ((Basset 1893), Shawiya (Penchoen 1973), Tarighit (Mettouchi, p.c.), Timimoum, Touat (Basset 1885), Harakta (Basset 1892), Chennini, Tamazret, Djerba (summarized in Brugnatelli, forthcoming), Sened (Provotelle 1911), Ghadames (Lanfry 1968; Kossmann 2013a, b), Zwara (Mitchell 1953), Sokna (Sarnelli 1924), Awjila (van Putten 2014), Tayert (Kossmann 2010), Tahaggart (Prasse 2010), Burkina Faso varieties (Sudlow 2011), Tamashek (Heath 2005), El Fogaha (Paradisi 1961), and Ghomara (Mourigh 2016)

Arabic influence on Berber varieties can arguably be found at many levels of these languages' grammar. While Arabic pronominal forms can appear with borrowed particles, nonintegrated verbs (Ghomara), and in certain syntactic contexts (Kossmann 2013a, b: 291–297), Berber independent personal pronouns are present everywhere. Few exceptions include the use of Arabic independent forms after the presentative *ha* in the Ayt Segrouchen (Morocco) variety (Kossmann 2017) or together with the Berber ones in Ghomara (Mourigh 2016).

Functions of Independent Personal Pronouns in Berber

Several studies have focused on pronominals affixes and clitics (Brugnatelli (1993, 1998), Kossmann (1997b, 2006), and Gutova (2017), among others), and others treat independent pronouns in relation to other pronominal forms (Galand 1966, 1994, 2010; Allaoua 1997; Anab 2000; Taine-Cheikh 2005; Aghali-Zakara 2006; Chaker 2015). Nevertheless, there are not many studies devoted to independent personal pronouns in Berber only, and many grammatical descriptions limit themselves to the list of these pronouns, together with those affixed to the verb. There are however a few exceptions, some of which are mentioned below.

van Putten (2014) lists the use of independent pronouns in Awjila Berber, for instance, when there is a “switch of subject in two consecutive sentences” (van Putten 2014: 108). Kossmann (2011: 78) remarks that they are used as a topicalized subject (1), or direct object (2), or focalized subject (3) in Ayer Tuareg:

- (1) *ənta* *t-əha* *ebəhun* *t-əybār* *d-əs*
 3S 3S:F-be.in:P EL:bag 3S:F-hide:LoP in-3s
 “She was hiding in the bag” (lit. she, she was in the bag, hiding it).
- (2) *năḳḳ* *dă* *wər_di_y-ewəd* *ăzāle*
 1S EMPH NEG_1S:ACC_3S:M-arrive:PN EA:song
 “Me, a song does not suffice me.”
- (3) *năḳḳ* *a* *y-əssán-ăn* *anḳen*
 1S NEUT PTC:S:M-know:LoP-PTC:S:M how

 a *he-dəs_təgi-m*
 NEUT NR_3S:DAT_2P:M-do:A-2P:M
 “I (am the one who) knows what you are going to do with her.”

Independent pronouns are often described in their use following a number of specific prepositions (viz., dative *i* “to,” Kossmann 1997a, b: 189–190; Kossmann 2011: 78; van Putten 2014: 107, or *am* “like,” Kossmann 1997a; Galand 2010: 110; van Putten 2014). Their use in nonverbal predication is also often mentioned (Kossmann 1997a: 188; Prasse 2010: 17; van Putten 2014: 108):

ənta tarabt “She is an Arab woman” (Prasse 2010: 17).

Independent pronouns are also used for pronoun coordination (you and me), usually with the preposition *did-*, *idd-* “with”:

- (4) *nək a=nɛr-əx=ku ká ir=a=nnə-ʃədd nək idd-ək ar=əlqədi*
I fut.=let.go:fut.-1s=DO.2sm neg. until=fut.=1p-go:fut. I with-2sm to=judge
“I will not let you go until you and I go to the judge” (van Putten 2014: 108).

and following the presentative or constructions such as “also; as well” (Kossmann 1997a: 189):

- (5) *ha netta*
voici lui
‘le voici’ [here he is].
- (6) *wala d netc*
même est moi
“moi aussi” [me too].

They are also commonly used in right and left dislocations (Kossmann 2012: 59) and often generally described as giving emphasis to the sentence (Prasse 1972: 180; Sudlow 2011: 68).

Independent Personal Pronouns in Siwi

Morphology and Syntax of Independent Pronouns

This section aims to provide a closer analysis of the use and function of independent personal pronouns in Siwi through examples taken from a corpus comprised of different kinds of topics and genres. Before analyzing the function of personal pronouns within the larger framework of the reference system of Siwi, and the pragmatic functions linked to the use of independent pronouns, the table below offers a glimpse on their morphology:

Siwi			
<i>Singular</i>			
	Independent Pronouns	Direct Object Pronouns	
			1SG
1	<i>niš</i>	<i>-i</i>	
2M	<i>šək</i>	<i>-ek</i>	<i>-šək</i>
2F	<i>šəm</i>	<i>-em</i>	<i>-šəm</i>
3M	<i>nətta</i>	<i>-t / -a</i> (Direct object pronouns have two possible forms. The one on the right is used after the verb stem (when the person indexes are prefixed only). The one on the left is used when it follows other suffixes (person indexes and indirect object pronouns).)	

(continued)

Siwi			
3F	əntatət	-tət / -et	
<i>PLURAL</i>			
1	ənšni (also ənčni, nišni)	-anax	
2	əknəm	-wən	-əknəm
3	əntnən	-tən / -en	

The table contains both independent personal pronouns and direct object clitics. They are presented together to show the similarities between these two series of pronouns (as opposed to other series, as mentioned earlier with reference to Galand's (1966) analysis in Section “[Forms of Independent Personal Pronouns](#)”) – but also because independent second-person pronouns are used when the verb bears the first-person singular index, as exemplified here:

- (7) ga-ɣɾ-á-šəm
IRR-see.AOR-1SG-DO.2SG.F
“(I will) see you!”

Another important feature of Siwi's independent personal pronouns (but not limited to them) is the absence of gender distinction in the plural (Souag 2013: 49–51), similar to what happens in Ghomara.

One of the most common uses of independent pronouns is with different types of nonverbal predications which in Siwi (unlike in many other Berber languages that use the particle *d* (Galand 2009)) is achieved simply through juxtaposition:

- (8) ənnhárdin / niš aħkík-a
in_the_past / IDP.1SG small.SG.M-PRAGM
“In the past, I was young...”
- (9) ənšní g šal n isíwan
IDP.1PL in town.SG.M of Siwa
“We are in the town of Siwa.”

As mentioned in Section “[Functions of Independent Personal Pronouns in Berber](#)” with regard to other Berber varieties, independent personal pronouns can appear after the preposition *i* “to.” The following example is a part of a common ending formula in folktales:

- (10) akəmpús n xer i ənšní / akəmpús n šar i əntnən /
bundle.SG.M of goodness to IDP.1PL / bundle.SG.M of badness to IDP.3PL /
“A bundle of goodness for us, a bundle of badness for them.”

Singular independent pronouns are used instead of same person plural pronouns when there is a need to specify the referents through *d* “and.”

In the following example, the speaker is talking about her daily routine. At lunch, she eats with her children, who have never been mentioned before. The first-person

3-go.PFV-PL 3-stay.PFV-PL in castle.SG.M far /
 yə-qqím-ən **nǽtta** **díd-əs** ssih
 3-stay.PFV-PL **IDP.3SG.M** **with-3SG** there
 “He took his daughter, they left, they went and stayed in a castle far
 away, they stayed together there.”

Second-person pronouns can be used for the vocative:

- (16) yə-m̄m-ás hey hey **šək** / yə-m̄m-ás ah
 3SG.M-say.PFV-IO.3SG hey hey **IDP.2SG.M** / 3SG.M-say.PFV-3SG.M yes
 He asked: “Hey, hey, you!” He replied: “Yes!”

And with imperatives:

- (17) tə-m̄m-ás háyya / **šək** óbdəd / gá-z̄r-aṭ
 3SG.F-say. come_on / **IDP.2SG.M** stand_up. / IRR-see.
 PFV-IO.3SG IMP AOR-2SG
 ““Come on, stand up, you will see!””

Third-person independent pronouns can be followed by full co-referent lexical nouns, in order to further precise who, or what, the referent is:

- (18) **nǽtta** **aggʷíd** yə-m̄m-ás /
IDP.3SG.M **man.SG.M** 3SG.M-say.PFV-IO.3SG /
 n wən g-yə-m̄má-y smiyət-én̄nəs /
 of REL IRR-3SG.M-say.AOR-IO.1SG name.SG.F-POSS.3SG /
 gá-š-ŋ-as lek̄r̄i /
 IRR-give.AOR-1SG-IO.3SG trophy.SG.M /
 “The man said: “The one who will tell me what is her name, I will
 give him a trophy!””
- (19) **əntátət** **tyər̄fət** ttǽlla abéhhər / af ágbən
IDP.3SG.F **room.SG.F** be_at.3SG.F north / on house.SG.M
 “The room is on the north side of the house.”

And independent pronoun can precede or follow *nnuba* “all”:

- (20) d i-tǽšsy-ən amán
 and 3-take.IPFV-PL water.PL
 i **əntn̄n** **nnúba** //
 to **IDP.3PL** **everybody** //
 “and they take from it water for everybody.”

Independent pronoun can also be the head of a cleft sentence.

In the following example, the speaker explains a game made in the sand, where obviously one of the participants is supposed to start. The predicate in the relative clause serves to identify or classify the pronominal head:

- (21) *niš əlbádi / yáʕni niš wən gá-bdu-x*
 IDP.1SG starter.SG.M / I_mean IDP.1SG REL IRR-start.AOR-1SG
 “I am the starter, I mean, I am the one who starts.”

Similarly, in a monologue, the speaker starts by saying that he wants to talk about “those who sing” in Siwa, and then he says that in Siwa, there is a group of people called “Zeggala.” With the following example, he establishes a relation between the Zeggala, taken up again with a 3PL independent pronoun, and those who sing, when they work in the garden:

- (22) *əntnən wən i-dəyyz-ən*
 IDP.3PL REL 3-sing.IPFV-PL
 “they are those who sing.”

Independent pronouns are also commonly used for fixed expressions like the following:

- (23) *niš smiyət-ənnaw*
 IDP.1SG name.SG.F-POSS.1SG
 “My name is...”
- (24) *af ʕəm*
 on IDP.2SG.F
 “It is up to you.”

Information Structure Functions

This section deals with the use of independent pronouns, word order, and information structure.

Word order, especially regarding the order of nouns in the preverbal or postverbal position, as well as the features of topic promotion (topicalization) and focalization, has been the subject of a number of studies on Berber (Galand 1964, 2010; Mettouchi 2003a, b; Lafkioui 2010, 2014; and Leguil 1986a, b: 115–120, for Siwi). The information structure functions of different word order constructions have been analyzed in other studies by Mettouchi for Kabyle (among which Mettouchi (2005), Mettouchi and Fleisch (2010) (Kabyle and Tashelhyt), and Mettouchi (2018)) or by Kossmann (2016) for Figuig Berber.

In addition, a comparative study between Kabyle (Algeria) and Siwi (Mettouchi and Schiattarella 2018) confirmed that different word orders had different

information structure functions. For the study, the verb, which is always marked by subject affixes (person indexes), was considered the reference point (Frajzyngier and Shay 2003). The V(O) (V, verb; O, direct object; S, subject, intransitive verb; A, agent, transitive verb; cl, direct object clitics) construction is considered in both languages as the default and described for its “topic continuation” function. All the other possible orders – when a noun co-referential with subject affixes (transitive or intransitive) or lexical direct objects (alone or co-referent with direct object affixes), as well as when both arguments are lexically expressed – were examined in the study, inside and outside the prosodic boundaries.

In Siwi, apart from the “topic continuation” function (V(O)), the following constructions and respective information structure functions have been described:

VS – VA_{cl}: new episode in narrative sequence (topicalization of an event or state), usually in a narrative sequence

SV – AV_{cl}: dependency (pragmatic or syntactic), used when a speaker gives an explanation, or comments regarding the previous intonation units.

When the two arguments of a transitive verb are lexically expressed, AVO is the only allowed order, thus having no single information structure function.

As for left and right dislocation, when the noun is before the prosodic boundary (+F0 rise), the function is called “subtopic shift”: from the one discussed in the previous intonation units to a different topic. On the contrary, when a noun is placed after the verb, in a separate intonation unit, its function is to reactivate a referent.

Despite the fact that lexical nouns and independent pronouns are different categories and have different syntactic functions, it is, nonetheless, useful to bear in mind the pragmatic functions mentioned above with regard to nouns; in fact, in the case of independent pronouns, we are dealing not just with third-person pronouns (which alternate with nouns and share their referential status) but also with first- and second-person pronouns. Although the latter are only used to indicate speech roles, under some constructions, their presence has specific information structure functions shared with third-person pronouns and nouns, as we will see in the section below.

Independent Pronouns and Word Order

(a) Independent Pronouns in Preverbal Position

When the independent pronoun precedes the verb, its function is similar to that of nouns’ (A or S) preceding verbs and personal affixes. This configuration is usually used to comment, summarize, or justify what has been said in the previous intonation units.

In the following example, a boy helps his father recover from his illness by buying him some medicines. His father cannot believe his son had the money to buy the medicine and asks him where he got it from. The son justifies himself, saying that he had been working:

- (25) way / səgmá kə́tr-aṭ wa /
 what / from_where bring.PFV-2SG DEM.SG.M /
 yə-ṛm-ás / niš xə́dm-ax / kətr-áx-tən //
 3SG.M-say.PFV-IO.3SG / IDP.1SG work.PFV-1SG / bring.PFV-1SG-DO.3PL //
 “‘What! From where did you get this?’ He replied: ‘I worked, I brought them.’”

(b) Independent Pronoun in the Postverbal Position

Independent pronouns following verbs are quite rare in our corpus. Their function is also analogous to the one found with nouns (S or A) following the verb. There is a “topicalization of an event or state,” presented as new (Mettouchi 2018; Mettouchi and Schiattarella 2018). The following intonation units develop this new topicalized event or state.

In the following example, the two protagonists challenge each other to see who will successfully prompt a couple to divorce. When one of the two (the devil) fails, the other (an old woman, the one who came up with the challenge) offers to show him how it is done. In fact, her success becomes the topic of the following intonation units:

- (26) tálti tláṣṣust tə-ṛm-ás gá-hh-ax niš
 woman.SG.F old.SG.F 3SG.F-say.PFV-IO.3SG IRR-go.AOR-1SG IDP.1SG
 “The old woman said: ‘I will go’.”

This construction should not be considered as strictly contrastive: the contrastive reading is probably influenced by the particular pragmatic context of this sentence (two participants challenging each other).

(c) Independent Pronouns Followed by a Verb and a Direct Object Noun

As mentioned in Section “[Information Structure Functions](#)”, there is no single information structure function when both the subject and the object of a transitive verb are present, because the order is fixed: the noun (or independent pronoun) precedes the verb and the object follows it.

In example (28), from a folktale, after a series of actions taken by the protagonist (a rooster), the storyteller makes the following comment, explaining how the rooster would manage to live with the girl (the most beautiful one) that he had just succeeded to get, through his ruse:

- (27) nótta yə-bnú tṣuṣṣət ánni
 IDP.3SG.M 3SG.M-build.PFV hut.SG.F COMP
 g-yə-gr-ən ləhyat-ənnəs əgd-əs
 IRR-3-live.AOR-PL life.SG.F-POSS.3SG in-3SG
 “He built a hut to live his life inside”

Contrastive Focus and Contrastive Topic

Independent pronouns are often found in contrastive focus or contrastive topic constructions.

In the corpus analyzed for this study, contrastive focus does not feature clefting, unlike other Berber languages (cf. Kabyle; Mettouchi 2003a, b; Galand 2010: 325–337, just to give an example).

Nevertheless, these constructions are characterized by a specific prosodic contour, where the F0 is higher right after the focused element (Schiattarella 2015).

In the following example, a goat is accused by a jackal of having dirtied the water. She replies, denying his accusations. In this case, the F0 is higher on the verb (*swi-t-a* “you drank”):

- (28) **šək** **swi-t-a** **uwwəl** /
 IDP.2SG.M **drink.PFV-2SG-PRAGM** first /
- šək** **swi-t-a** **zdát** /
 IDP.2SG.M **drink.PFV-2SG-PRAGM** in_front_of /
 “YOU drank first, YOU drank before (me).”

With contrastive topics, on the contrary, this idea of correction, or contradiction, characteristic of contrastive focus, is absent (Lambrecht 1994: 291). The syntax of contrastive topics and focus is the same, but the prosodic contour is different. For contrastive topics, the F0 is stable (Schiattarella 2015):

- (29) **šəm** **ga-γəřs-at** **imán-nəm** /
 IDP.2SG.F **IRR- slaughter.AOR-2SG** oneself-POSS.2SG.F /
- niš** **g-ús-ix** / **ga-səmm-á-šəm**
 IDP.1SG **IRR-come.AOR-1SG** / **IRR-cook.AOR-1SG-DO.2SG.F**
 “You will slaughter yourself, I will come and I will cook you.”
- (30) **əntátət** **tə-šmār** **taməllált** /
 IDP.3SG.F **3SG.F-do.PFV** white.SG.F /
- wəłtmá-tsən** **tə-šmār** **tazəttáft**
 sister.SG.F-POSS.3PL **3SG.F-do.PFV** black.SG.F
 “She became white, their sister became black.”

There are also cases where independent pronouns are not explicitly contrasted to other referents, as is the case in the previous examples where we have a binary opposition between “you vs. me” in (31) or “she vs. their sister” in (32). Nevertheless, the contrast is still present, and in some cases, independent personal pronouns are co-referent with possessive pronouns, reinforcing the idea of contrast. In (33), seven brothers do not believe that their sister has black skin (because of a previous episode where a slave made her bathe in a source of water that darkened her).

The fact that they insist on her sister having white skin is in contrast with the way she presents herself (with black skin), despite the fact that she is indeed their sister:

- (31) níš wəltmá-twən / yə-m̄m-án-as ulá /
 IDP.1SG sister.SG.F-POSS.2PL / 3-say.PFV-PL-IO.3SG no /
 mólmi wəltmá-tnax tazəttáft / ənšní wəltmá-tnax taməllált
 when sister.SG.F-POSS.1PL black.SG.F / IDP.1PL sister.SG.F-POSS.1PL white.SG.F
 “‘I am your sister!’ (The brothers) said (to the real sister): ‘No, how come our sister is black? Our sister is white!’”

Focus Particles

Independent pronouns are also used with a series of additive, scalar, or restrictive focus particles (also, even, only):

bídu “also” following the focalized pronoun:

- (32) ámra níš bídu fəkkɾ-áx-a
 now IDP.1SG also remember.PFV-1SG-PRAGM
 “Now, I remember too.”
- (33) fá y-uṭən nətta bídu
 so 3SG.M-be_sick.PFV IDP.3SG.M also
 “So, he got sick too.”

ħətta “even,” preceding the focalized pronoun:

- (34) yə-ɣɾ-a ħətta nətta
 3SG.M-see.PFV-DO.3SG.M even IDP.3SG.M
 “Even him saw it”

yɛr “only, except” preceding the focalized pronoun:

- (35) t́-ɖdər yɛr əntátət
 3SG.F-live.PFV only IDP.3SG.F
 “‘She was the only one to survive.’”

The restrictive particle *bass* “only” follows the focalized element:

- (36) əntnən bass
 IDP.3PL only
 “Only them.”

The latter two restrictive focus particle *bass* and *yɛr* also function in Siwi as adversative conjunctions or as discourse markers and are often found in left dislocation constructions (preceding the pronoun or the noun phrase):

- (37) tašəbbabt / t-šəmmar am zzmar /
 tashebbabt.SG.F / 3SG.F-do.IPFV like zzmar.SG.M /
bass əntátət / t-šəmmar s ləhdid /
but IDP.3SG.F / 3SG.F-do.IPFV with iron.SG.M
 “The tashebbabt (flute) is like the zzmar, but it (tashebbabt), it is made with iron.”
- (38) niš ga-ššáhh-ax / ga-ktər-š-áwən wən xs-əm /
 IDP.1SG IRR-heal.AOR-1SG / IRR-bring.AOR-1SG-IO.2PL REL want.PFV-2PL /
yer šək / kəmməl timədrast-ənnək
but IDP.2SG.M / continue.IMP school.SG.F-POSS.2SG.M
 “I will be better, I will bring you what you need, but you, finish your school!”

Conclusions

This paper has examined the system of independent personal pronouns in Berber, highlighting their typological features. It has also introduced the syntactic and pragmatic functions of these pronouns, with a brief overview of some Berber languages and a more extensive discussion for Siwi Berber through the analysis of different texts, of different genres and topics.

Independent pronouns have often been analyzed in the literature with a focus on the features linked to their origin or source of grammaticalization, their referential role (especially the different reference structures of first- and second-person pronouns and third-person ones), and the semantics of gender and number vis-à-vis nouns and their functions, especially when compared to other means of reference used in a given language.

In Berber, independent pronouns are often associated with direct object pronouns (even though this series of pronouns often has two paradigms and is not only used for direct objects) because of their formal similarities (Galand 1966). They mark person, gender, and number through different affixes. When present, gender affixes are the same for singular and plural, and number affixes are the same for all persons. Internal developments are, nevertheless, very common, and each language has to be regarded on its own. The majority of Berber languages mark gender distinction for second- and third-person singular and non-singular. Many Berber languages, additionally, mark gender distinction at the first-person plural. Among the different paradigms attested in Berber, Tuareg varieties differ from all others because they mark gender distinction in the plural, for all persons, but in the singular only for second-person pronouns. Ghomara in Morocco and Siwi are among the few Berber languages that do not mark gender distinction in the plural. The lack of gender distinction in the plural is quite consistent in Siwi and not limited to these pronouns. This is probably a consequence of the influence through contact of some sedentary

Arabic varieties (Souag 2013: 51). Finally, Berber languages do not make any distinction between inclusive and exclusive first-person plural pronouns.

As a case study, the position of the Siwi independent pronouns in verbal clauses has been investigated. Each order has its specific information structure function, similar to the ones found when a lexical noun co-referent to the person indexes is present.

Moreover, independent pronouns are often used to express contrast between two referents (contrastive topic), or contradiction, or correction (contrastive focus). Nevertheless, the use of independent pronouns is not automatically linked to the idea of contrast or emphasis. Personal pronouns can also be found in left-detached constructions or with a number of focus particles. Right-detached pronouns alone are not found in my corpus, but this aspect rates a further discussion.

Acknowledgement I would like to thank the Alexander von Humboldt Foundation for supporting my research. I am also grateful to the editors of the volume, the reviewers of the paper for their comments and suggestions, and the Siwi speakers for sharing their knowledge with me

References

- Aghali-Zakara M (2006) Pronoms berbères – unité et diversité du système touareg. *Études berbères* III – Le nom, le pronom et autres articles. Actes du « 3. Bayreuth-Frankfurter Kolloquium zur Berberologie », Bayreuth, 1–3 juillet 2004. Köln: Rüdiger Köppe Verlag [Berber Studies 4]
- Allaoua A (1997) Sur Les Pronoms Personnels: Questions D'autonomie Primitive. *Études et Documents berbères* 15:105–117
- Anab A (2000) Études morphosyntaxique des pronoms personnels en berbère (Dialecte tachelhiyt d'Agadir et sa région). *Études et Documents Berbères* 18:185–203
- Basset R (1885) Notes de lexicographie berbère. [Quatrième série], [Vocabulaire du Touat et du Gourara, argot du Mzab, dialecte des Touaregs Aouelimiden]. Imprimerie Nationale, Paris
- Basset R (1892) Notice sur les dialectes berbères des harakta et du Djerid tunisien [Publications du 9ème Congrès International des Orientalistes, Londres 1891]. Oriental University Institute, London
- Basset A (1952) La langue berbère. Oxford University Press, London
- Basset R (1893) Étude sur la Zenatia du Mzab, de Ouargla et de l'Oued-Rir'. Leroux, Paris
- Beguino F (1921) Gli studi berberi dal 1919 al maggio 1922. *Rivista degli Studi Orientali*, IX, pp 382–408
- Beguino F (1942) Il Berbero Nefûsi di Fassâto: Grammatica, testi raccolti dalla viva voce. Vocabolarietti, 2nd edn. Istituto per l'Oriente, Rome
- Benrejda N (2015) Étude de géographie linguistique berbère: cas des pronoms personnels autonomes. Master thesis, University Mouloud Mammeri, Tizi Ouzou (Algeria)
- Benveniste E (1947) Structure des relations de personne dans le verbe. *Bulletin de la Société de Linguistique de Paris* 43(1):126
- Bhat SDN (2004) Pronouns. Oxford University Press, Oxford
- Bhat SDN (2013) Third person pronouns and demonstratives. In: Dryer M, Haspelmath M (eds) The world atlas of language structures online. Max Planck Institute for Evolutionary Anthropology, Leipzig. <https://doi.org/10.5281/zenodo.7385533> (Available online at <http://wals.info/chapter/43>, Accessed 17 Apr 2023)
- Brugnatelli V (1993) Quelques particularités des pronoms en berbère du Nord. In: Drouin J, Roth A (eds) A la croisée des études libyco-berbères. Mélanges offerts à Paulette Galand-Pernet et Lionel Galand. Geuthner, Paris, pp 229–245

- Brugnatelli V (1998) Encore à propos des pronoms berbères. *Comptes Rendus du Groupe Linguistique d'Études Chamito-Sémitiques* 32(1988–1994):151–158
- Brugnatelli V (forthcoming) Tunisie (berbère). In: Chaker S (ed) *Encyclopédie berbère*. Peeters, Paris/Leuven
- Chaker S (2004) Les paradigmes personnels du berbère. In: Ibriszimow D, Segerer G (eds) *Systèmes de marques personnelles en Afrique*. Peeters, Paris/Leuven, pp 43–54
- Chaker S (2015) Pronoms. In: Chaker S (ed) *Encyclopédie berbère*. Paris and Leuven, Peeters
- Cysouw M (2013) Inclusive/exclusive distinction in independent pronouns. In: Dryer M, Haspelmath M (eds) *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology, Leipzig. <https://doi.org/10.5281/zenodo.7385533> (Available online at <http://wals.info/chapter/39>, Accessed 17 Apr 2023)
- Daniel M (2013) Plurality in independent personal pronouns. In: Dryer M, Haspelmath M (eds) *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology, Leipzig. <https://doi.org/10.5281/zenodo.7385533> (Available online at <http://wals.info/chapter/35>, Accessed 17 Apr 2023)
- Destaing E (1907) Étude sur le dialecte berbère des Beni-Snous. Leroux, Paris
- Frajzyngier Z, Shay E (2003) Explaining language structure through systems interaction, *Typological studies in language* 55. John Benjamins, Amsterdam/Philadelphia
- Gabsi Z (2003) An outline of the Shilha (Berber) Vernacular of Douiret (South Tunisia). PhD dissertation, University of Western Sydney (Australia)
- Galand L (1964) L'énoncé verbal en berbère. Étude de fonctions. *Cahiers Ferdinand de Saussure* 21: 33–53
- Galand L (1966) Les pronoms personnels en berbère. *Bulletin de la Société de linguistique de Paris* 61(1):286–298
- Galand L (1994) La personne grammaticale en berbère. *Faits de Langues* 3(1):79–86
- Galand L (2009) *La particule prédicative* d. In: Chaker S, Mettouchi A, Philippson G (eds) *Études de phonétique et linguistique berbères. Hommage à Naïma Louali (1961–2005)*. Peeters, Paris and Leuven, pp 141–154
- Galand L (2010) Regards sur le berbère. *Centro Studi Camito-Semitici di Milano*, Milan
- Gutova E (2017) Direct object clitic pronouns in Ketama/Senhaja Berber (Northwestern Morocco). *Études et Documents Berbères* 37:105–148
- Gutova E (2021) *Senhaja Berber varieties: phonology, morphology, and morphosyntax*. PhD thesis, Université Sorbonne Nouvelle, Paris
- Haspelmath M (1997) *Indefinite Pronoun*, Cambridge University Press
- Heath J (2005) *A grammar of Tamashek (Tuareg of Mali)*. Mouton, Berlin
- Heine B, Song K-A (2010) On the genesis of personal pronouns: some conceptual sources. *Lang Cogn* 2(1):117–147
- Heine B, Song K-A (2011) On the grammaticalization of personal pronouns. *Journal of Linguistics* 47(3):587–630
- Helmbrecht J (2002) *Personal pronouns – form, function and use*. University of Erfurt Habilitationsschrift, Erfurt
- Helmbrecht J (2013) Politeness distinctions in pronouns. In: Dryer M, Haspelmath M (eds) *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology, Leipzig. <https://doi.org/10.5281/zenodo.7385533> (Available online at <http://wals.info/chapter/45>, Accessed 17 Apr 2023)
- Kossmann M (2016) On word order in Figuig Berber narratives: The uses of pre- and postverbal lexical subjects. *Wiener Zeitschrift für die Kunde des Morgenlandes*, 106:115–146
- Kossmann MG (1997a) *Grammaire du parler berbère de Figuig*. Peeters, Paris and Leuven
- Kossmann MG (1997b) Le pronom d'objet direct de la troisième personne en berbère. In: Tosco M, Bausi A (eds) *Afroasiatica Neapolitana, Studi Africanistici*. Istituto universitario orientale, Naples, pp 69–79
- Kossmann MG (1999) *Essai sur la phonologie du Proto-Berbère*. Rüdiger Köppe Verlag, Köln
- Kossmann MG (2000) *Esquisse Grammaticale Du Rifain Oriental*. Peeters, Paris and Leuven

- Kossmann MG (2006) La reconstitution des pronoms satellites de la deuxième personne du pluriel en berbère. In: Allati A (ed) *Linguistique amazighe: les nouveaux horizons*. Université Abdelmalek Essaâdi, Faculté des Lettres et des Sciences Humaines, Tétouan, pp 28–47
- Kossmann MG (2011) *A grammar of Ayer Tuareg (Niger)*, Berber Studies 30. Rüdiger Köppe Verlag, Köln
- Kossmann MG (2012) Berber. In: Frajzyngier Z, Shay E (eds) *The Afroasiatic languages*. Cambridge University Press, Cambridge, pp 18–101
- Kossmann MG (2013a) The Arabic influence on northern Berber, *Studies in Semitic languages and linguistics*. Brill, Leiden
- Kossmann MG (2013b) *A grammatical sketch of Ghadames Berber (Libya)*, Berber studies 40. Rüdiger Köppe Verlag, Köln
- Kossmann MG (2017) Personal pronouns in the Ayt Seghrushen Berber variety of the province of Taza. In: Allati A (ed) *Mélanges Michael Peyron*, Berber Studies 47. Rüdiger Köppe Verlag, Köln, pp 253–264
- Kossmann MG (2020) Berber subclassification. In: Vossen R (ed) *The Oxford handbook of African languages*. Oxford University Press, Oxford, pp 281–289
- Lafkioui M (2007) *Atlas linguistique des variétés berbères du Rif*, Berber Studies 16. Rüdiger Köppe Verlag, Köln
- Lafkioui M (2010) La topicalisation en berbère: Formes et structures. In: Ibrizimow D, Vossen R, Stroemer H, Kossmann M (eds) *Studien zur Berberologie*, Berber Studies 5. Rüdiger Köppe Verlag, Köln, pp 121–132
- Lafkioui M (2014) Topicalization in Berber: a typological perspective. In: Taine-Cheikh C, Lux C (eds) *Berber in typological perspective [A special issue of language typology and language universals]*. *STUF* 67(1):97–112
- Lafkioui M (2021) The pronoun in Rif Berber (from Senhaja to Iznasen). In: Merolla D, Cauber D, Nait Zerad K, Cassuto P (eds) *Les études berbères à l'ère de l'institutionnalisation de tamaziɣt. Mélanges en l'honneur de Salem Chaker et Abdellah Bounfour*. L'Harmattan, Paris, pp 197–225
- Lambrecht K (1994) *Information structure and sentence form: topic, focus and the mental representations of discourse referents*. Cambridge University Press, Cambridge
- Lanfray J (1968) *Ghadamès I. Étude linguistique et ethnographique*. Fichier de Documentation Berbère, Fort-National
- Laoust E (1912) *Étude sur le dialecte berbère de Chenoua. Comparé avec ceux des Beni Menacer et des Beni Salah*. Leroux, Paris
- Leguil A (1986a) Notes sur le parler berbère de Siwa (II). *Bulletin des Etudes Africaines de L'Inalco* VI(12):97–124
- Leguil A (1986b) Notes sur le parler berbère de Siwa (II). *Bulletin des Etudes Africaines de L'Inalco* VI(12):97–124
- Lux C (2013) *Le Tetsrerret, Langue Berbère Du Niger*, Berber Studies 38. Rüdiger Köppe Verlag, Köln
- Mettouchi A (2003a) Contrastive focalization on clefts in Taqbaylit Berber. In: Mettouchi A, Ferré C (eds) *Proceedings of IP2003 interfaces Prosodiques*, Nantes 27–29 mars, pp 143–148. <https://llacan.cnrs.fr/pers/mettouchi/ip2003/actes-ip2003.pdf>
- Mettouchi A (2003b) Focalisation contrastive et structure de l'information en kabyle (berbère). In: François J, Lacheret A (eds) *Mémoires de la Société de Linguistique de Paris : Fonctions et moyens d'expression de la focalisation*. Peeters, Paris-Leuven, pp 81–97
- Mettouchi A (2005) Discourse-configurationality and the encoding of semantic macroroles in Taqbaylit Berber. In: Mengozzi A (ed) *Studi afroasiatici. Contributi presentati all'XI Incontro italiano di linguistica camito-semitica* (Bergamo, 5–7 giugno 2003), *Materiali linguistici* 52. Franco Angeli, Milano, pp 83–96
- Mettouchi A (2018) The interaction of state, prosody and linear order in Kabyle (Berber): grammatical relations and information structure. In: Tosco M (ed) *Afroasiatic: data and perspectives*. John Benjamins, Amsterdam and Philadelphia, pp 261–285

- Mettouchi A, Fleisch A (2010) Topic-focus articulation in Taqbaylit and Tashelhit Berber. In: Fiedler I, Schwarz A (eds) *The expression of information structure: a documentation of its diversity across Africa*, ZAS Berlin, series typological studies in language 91. John Benjamins, Amsterdam/Philadelphia, pp 193–232
- Mettouchi A, Schiattarella V (2018) The influence of the state distinction on word order and information structure in Kabyle and Siwi (Berber). In: Adamou E, Haude K, Vanhove M (eds) *Information structure in lesser-described languages: studies in prosody and syntax*. John Benjamins, Amsterdam/Philadelphia, pp 265–296
- Mitchell TF (1953) Particle-noun complexes in a Berber dialect (Zuara). *Bull Sch Orient Afr Stud* 15(2):375–390
- Mithun M (1990) Third-person reference and the function of pronouns in central Pomo natural speech. *Int J Am Linguis* 56(3):361–376
- Mourigh K (2016) *A grammar of Ghomara Berber (North-West Morocco)*, Berber Studies 45. Rüdiger Köppe Verlag, Köln
- Naumann C (2012) Acoustically based phonemics of Siwi (Berber), Berber Studies 30. Rüdiger Köppe Verlag, Köln
- Nehil M (1909) *Étude sur le dialecte de Ghat*. Publications de l'École des lettres d'Alger. Leroux, Paris
- Paradisi U (1961) El-Fôgâha, Oasi Berberofona del Fezzân. *Rivista degli Studi Orientali* 36: 293–302
- Penchoen TG (1973) *Étude syntaxique d'un parler berbère*. Centro di Studi Magrebini, Napoli
- Prasse K-G (1972) *Manuel de grammaire touarègue (tāhāggart): Phonétique. Écriture. Pronom*. Edition de l'Université de Copenhague, Copenhague
- Prasse K-G (2010) *Tuareg elementary course (Tahāggart)*, Berber Studies 29. Rüdiger Köppe Verlag, Köln
- Provotelle R (1911) *Étude sur la tamazir't ou zénatia de Qalaât es-Sened (Tunisie)*. Leroux, Paris
- Renisio A (1932) *Étude sur les dialectes berbères des Beni Iznassen, du Rif, et des Senhaja de Sraïr; grammaire, textes et lexique*. Publications de l'Institut des hautes études marocaines. Leroux, Paris
- Samelli T (1924) *Il dialetto berbero di Sokna: Materiali lessicali, testi manoscritti in caratteri arabi, con trascrizione e traduzione*. Supplemento all'Africa Italiana, VI, pp 1–46
- Schiattarella V (2015) *Le berbère de Siwa: documentation, syntaxe et sémantique*. PhD dissertation, Ecole Pratique des Hautes Etudes, Paris
- Schiattarella V (2017) *Berber texts from Siwa: including a grammatical sketch; with a foreword by Lionel Galand*, Berber Studies 46. Rüdiger Köppe Verlag, Köln
- Siewierska A (2013) Gender distinctions in independent personal pronouns. In: Dryer M, Haspelmath M (eds) *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology, Leipzig. <https://doi.org/10.5281/zenodo.7385533> (Available online at <http://wals.info/chapter/44>, Accessed 17 Apr 2023)
- Souag L (2013) *Berber and Arabic in Siwa (Egypt): a study in linguistic contact*, Berber Studies 37. Rüdiger Köppe Verlag, Köln
- Sudlow D (2011) *The Tamasheq of north-East Burkina Faso: notes on grammar and syntax including a key vocabulary*, Berber studies 1. Rüdiger Köppe Verlag, Köln
- Taïfi M (1991) *Dictionnaire Tamazight-Français*. L'Harmattan-Awal, Paris
- Taine-Cheikh C (2005) Les marques de 1ère personne en berbère. Réflexions à partir des données du zénaga. In: Mengozzi A (ed) *Studi Afroasiatici. XI Incontro Italiano di Linguistica Camitosemitica*. Franco Angeli, Milano, pp 97–112
- van Putten M (2014) *A grammar of Awjila Berber (Libya). Based on Umberto Paradisi's material*, Berber Studies 41. Rüdiger Köppe Verlag, Köln
- Vycichl W (2005) *Berberstudien and a sketch of Siwi Berber*, Berber Studies 10. Rüdiger Köppe Verlag, Köln

Action Nouns in Zenaga Berber of Mauritania

30

Catherine Taine-Cheikh

Contents

Introduction	638
Vowel Patterns of Action Nouns	639
Bisyllabic ANs in “Non- <i>a</i> ”	640
Bisyllabic ANs in “ <i>a</i> ”	640
AN in “ <i>a</i> —Non- <i>a</i> ”	641
AN in “Non- <i>a</i> — <i>a</i> ”	642
Plurisyllabic AN in “Non- <i>a</i> ”	643
Plurisyllabic AN in “ <i>a</i> ”	643
AN in “ <i>a</i> — <i>a</i> —Non- <i>a</i> ”	644
AN in “ <i>a</i> —Non- <i>a</i> —Non- <i>a</i> [—Non- <i>a</i>]”	645
Other Cases	646
Summary Table of the Main Patterns	648
Analysis and Comments	649
From the Point of View of the Root	649
From a Morphological Point of View	652
Conclusion	655
Bibliography	656

Abbreviations

A	Aorist
AN	Action noun
FAN	Feminine action noun
Fq	Frequency
IA	Intensive aorist (or imperfective)

C. Taine-Cheikh (✉)

LACITO (CNRS, Université Sorbonne Nouvelle et Inalco), Paris, France

e-mail: cath.tainecheikh@gmail.com

MAN	Masculine action noun
NP	Negative perfective
P	(Positive) perfective
U[F]AN	Unitary [feminine] action noun

Introduction

As in many languages, lexemes in Berber are divided into several syntactic classes. The two most important are verbs and nominals. The case of adjectives is more complex: in some varieties of Berber, adjectives behave like a subclass of nominals, while in others they constitute a particular syntactic class, intermediate between the two main classes (Taine-Cheikh 2014). Zenaga belongs to the second group, and this is the reason why this study opts to engage conjugated adjectives rather than conjugation of qualitative verbs, as not all verbs with qualitative meaning have a conjugation distinct from that of other verbs.

Many verbs have nouns with more or less predictable meaning that can be called verbal nouns. Some of them, with a nasal prefix (usually *m*, more rarely *n*), mainly provide agent or habit nouns. Others, with a sibilant (or “*S*”) prefix, are the source, among others, of instrumental nouns. The most numerous, however, have intrinsic semantics linked to that of the conjugated form and mean “the fact of. . .” Although they concern action verbs as well as stative verbs and even, in the case of Zenaga, conjugated adjectives, here they are called action nouns (henceforth AN) to distinguish them from other verbal nouns.

Galand (2002) and Taifi (2006) have precisely described the syntactic uses of ANs and in particular the “internal complement” constructions where the complement is a nominal derivative (AN) of the same root as the verb. The anteposition of the AN, as in *tarəwla irwəl* “the flight, he fled,” and *tabəxni ibəxxin* “the whiteness, he is white” (Taifi 2006: 245), is impossible in Zenaga. On the other hand, Zenaga allows the more frequent construction where the internal complement, postposed to the verb (transitive or not), itself bears a determiner, for example, *āmūrih yuḍṛaḇ mārih aḍṛuḇ aḇuḇan* “the employee seasoned the rice well” (lit. “. . .with a good seasoning”). On the other hand, Zenaga uses ANs to express a high degree, in an exclamatory construction of the type *əḍ + AN + ən. . .* which is specific to this language (Taine-Cheikh 2012: 280–283), for example, *əḍ əymanḍr-ən-k!* “how beautiful you are!” (lit. “with beauty-of-yourself”).

While ANs seem to be used in almost all varieties of Berber – their presence is sometimes quite discrete, as in Ghomara where the majority of ANs come from Arabic (Mourigh 2016: 123), a detailed description is not often available. It is proposed here to complete the study started in 2012 on the formation of ANs in Zenaga.

In the first part (2.), the different vowel patterns found are inventoried in relation to the gender of the AN and the properties of the verbs (semantics and/or bi- vs. plurisyllabic form). In the next part (3.), all the characteristics of the AN are

analyzed, particularly from the point of view of the root (gemination vs. de-gemination, realization of certain radicals and of the feminine suffix) and verbal patterns. Given the complexity of the phonological system in synchrony – due to the regular and irregular effects of diachronic evolution – it should be noted that in Zenaga, the correspondences between single and geminate consonants are only partly predictable – the realizations of simple consonants are generally more “relaxed,” except in some particular contexts (Taine-Cheikh 2003). The most frequent correspondences are *dd/ḏ* (for [ḏ]), *tt/t* and *ḏ, ḏd/ḏ* (for [ḏ]), *zz/z* (for [θ]) and *ẓ* (for [ʒ]), *zz/ẓ* (for [θ]), *ẓẓ* (for [ʒ]) *ẓ̣* (for [ʒ]), *ss/ṣ* and *šš* (for [ʃ] et [ʃʃ]), and *qq/q̣*. Geminates are realized as tensed consonant (noted by a capital letter) when they are in coda – especially in finals when the *-t* of the feminine singular is assimilated to the last radical (cases of *T, D, Ḑ, Z, Ẓ, and S*).

Vowel Patterns of Action Nouns

In Zenaga, some basic principles underlie the functioning of the morphological system – especially that of verbs (see Cohen and Taine-Cheikh (2000) and Taine-Cheikh (2009)).

- (i) The vowel system has three short phonemes (*a*, *i*, and *u* —with a frequent reduction of the /i/ vs. /u/ opposition to /ə/) and three long phonemes (*ā*, *ī*, and *ū*), but the verbal morphology is based on an interplay of oppositions between “*a*”-type vowels (*a* and *ā*) and “non-*a*”-type vowels (*i*, *u*, *ə*, *ī*, and *ū*).
- (ii) The aorist (A) and imperative (I) modes and the indicative aspects, perfective (P) vs. imperfective (IA) – positive vs. negative – of the verb are mainly determined by the vowel patterns, but a distinction must be made between basic (“bare”) and derived (prefixed) verbs and between bisyllabic and trisyllabic (or even quadrisyllabic) verbs.
- (iii) While passive forms of active verbs are derived by prefix, like causative and reflexive verbs, there are middle (or internal) verbs and stative verbs, which constitute, by virtue of their vowel patterns, particular verbal subcategories.

Since every non-passive verb (action, stative or internal) has a potential corresponding AN in Berber, it is questionable whether these three main principles have the same importance for this area of nominal morphology; however, there is still the distinction between “*a*” and “non-*a*” vowels.

What follows first engages with the most frequent patterns and groups together the rarest section “Other Cases”. For each vowel pattern, it will look at the masculine AN (MAN) and then the feminine AN (FAN), generally less frequent.

Unless otherwise specified by (P) for perfective, the verbs will be given here in an imperative form (I) or, failing that, with the aorist pattern (A). The meaning of the AN, unless otherwise specified, corresponds to “fact of. . .” or “state of. . .”

There are many variants, often without any semantic difference (including between MAN and FAN), but when there is a difference, it is indicated – in particular by the use of UAN if one of the variants has the meaning of a unitary noun.

Bisyllabic ANs in “Non-a”

Masculine ANs

A large number of bisyllabic verbs have a “non-a” MAN. In this case, the AN has the same pattern as the negative perfective (which, for some verbs, is identical to the aorist). In a few rare ANs, one of the “non-a” vowels (or even both) is an *a*, for example, AN *əgəf* from *əgəf* “be afraid”; AN *ugməš* from *əgməš* “make a knot; be miserly.” However, the most common “non-a” vowel is *u* (or *ū*): AN *ūnnəg* from *ūnnəg* “be afflicted with ophthalmia”; AN *unkur* from *ānkur* “stand up; explode (gun)”; *unkur* “be scraped (bone)”; AN *užun* from the stative verb *ažun* “be shared” and from the action verbs *užun* “share (oneself),” and *ažun* “share.” These examples show that this pattern can be that of an action or stative verb. *ərīh* “be raw” is, however, the only case of an AN in “non-a” corresponding to a conjugated adjective, that of *ārāh* “raw.”

Feminine ANs

FANs are quite frequent (about 30) but much less frequent than MANs of the same pattern. In some cases, they coexist with MANs, regardless of whether they have the same pattern. They can then be simple variants of the MAN, for example, AN *təttəL* from *ətti* “be rolled up; roll up,” variant of *ətti*. They may have a more specialized meaning, for example, AN *tuzziʔd* from *azzi* “plait hair,” while *uzzi* is also an AN from *azzi* in the sense of “weave (a mat).” Finally, they can have a clear sense of a one-time (or unitary) noun, for example, UAN *tīkih* (/AN *ukfi*) from *ākfi* “give” and UAN *tuʔm̄muḏ* (/AN *uʔmuḏ*) from *uʔmuḏ* “close of itself.”

In general, however, FANs are the only ones used, possibly with a unitary noun meaning, thus AN *tugguZ* from *yoʔgəž* “know”, AN *tum̄muZ* from *aʔmuž* “bite,” AN *tīrih/tiʔrih* from *īrih/iʔrih* “dictate,” AN *təd(d)rih* from *āddər* “sting,” AN *tiʔgərt/tiʔgrih* from *oʔgər* “fly,” UAN *tuʔšbih* from *aʔžuf* “take a sip, a puff,” and (U)AN *tīngəmt* from *īngəm* “have a craving (for a pregnant woman).”

In three cases, the verb is monosyllabic, thus AN *tīzih* from *ūz* “be near parturition.”

Bisyllabic ANs in “a”

Masculine ANs

āwāy “speech, language” serves as the AN of the derived verb (*iš*)*šīwiy* “speak” but MANs in “a—a” are usually the ANs of bisyllabic verbs. These ANs may be the only variant in use (five cases): AN *aʔəf(f)* from *ārəf(f)* “cook in water”; AN *āyāš* from *āyəs* “mow”; AN *aʔyāš* from *ayūs* “clean”; AN *āybbād* from *ābḥud* “jump; leap”; AN *aʔār* from *aʔār* “play.” But these are more often variants of an AN in

“non-*a*—non-*a*.” Sometimes the meaning differs, such as in AN *älläg* from *ällig* “become calm; stand still” and AN *əlləg* from *ällig* “be calm.” The meaning is more specific for AN *änäṣ* from *ānuṣ* “remove liquid from the pot” than for AN *unuṣ* from *ānuṣ* “decrease.” It is less so for *aṣṣa* “braiding hair; weaving (a mat)” than for *uṣṣi* and *tuzziḍ*, other ANs from *aṣṣi*. More frequently (11 cases), the meaning is the same, for example, AN *aynās/uyniṣ* from “put on the fibulae,” AN *aynās/ugnaṣ* from *ognāṣ* “argue,” AN *aḥraḍ/ufṣruḍ* from *āḥruḍ* “oblige,” AN *āyraḍ/ugruḍ* from *āgruḍ* “scrape (to clean up),” AN *ārāṣ/əraṣ* from *āriṣ* “be late or make late,” and *ättäy* variant from *əttiṣ* (and *təttəL*).

Feminine ANs

The FAN in “*a*—*a*” is frequent. First of all, it is almost the rule for conjugated adjectives, whether they express a color, a defect, or some other qualities. Examples for it are: *tbäyḍayḍ* “being blue or green (*bäyḍig*),” *tbaṣbaḍ* “being of two colors” other than white and black (*baṣbuḍ*),” *tkäbbānt* “state of being deep (*käbbun*),” *tkägās* “being blind (*käygiṣ*),” *tšäyḥ(f)aḍ* “being unsightly (*šäyḥ*),” *taʔṣzagt* “being heavy (*aʔzag*),” *täṣaʔrt* “being thick, fat (*ṣaʔw(w)ur*),” *tkäyāl* “being short or dwarfed (*käyiy*),” and *taṣṣaḍ* “being stinky (*raṣṣuḍ*).”

This FAN is very often used for both the conjugated adjective and the corresponding verb expressing becoming (in this case, the latter is most often a trilateral verb), for example, *tädāl* “be (*ädäy*) or become (*idiy*) black,” *tkaʔräs* “be (*koʔriṣ*) or become (*ukkuʔriṣ*) active,” *ṭhaṣḥäft* “be (*ḥaṣḥuf*) or become (*uḥḥurḥuf*) hard,” and *tgäyḡamt* “be (*gäyḡum*) or become (*uggiyḡam*) soft, smooth.”

Many other verbs have this same FAN, often as the only AN, for example, AN *taḍah* from *aḍih* “pant,” AN *taʔškäkt* from *aʔṣi* “build,” AN *ṭayäṭ* from *aṣyih* “wait,” AN *tännäḡt* from *innəg* “avoid,” AN *täyṣäkt* from *okṣi* “graze” and *äṭʔši* “dress,” AN *tkäsräh* from *okṣər* “go downstairs,” AN *tmäṣṣäḍ* from *änṣih* “sell,” AN *tärdäh* from *äräd* “bathe,” and AN *träwyä* from *ärwiyi* “flee.” This is often also a variant of a MAN (sometimes in “*a*,” but more often in “non-*a*”) like AN *tädḡräh* (*/ädḡər*) from *ädḡər* “be decorated; decorate,” AN *tädämnäh* (*/ädḡmən*) from *ädḡmən* “let flour stand with yeast (so that it rises),” AN *tgämkäh* (*/ugḡmuk*) from *ägḡmug* “follow,” AN *tgänäh* (*/ugun*) from *ägun* “barrack,” AN *tkäṣṣäft* (*/ukṣuf*) from *äkṣuf* “shame,” and AN *täṣyäh* (*/əṣiy*) from *äṣiy* “be quick.”

Note that a number of trisyllabic verbs have a bisyllabic FAN, such as AN *tgaʔfäṭ* from *ägguʔfəh* “go north,” AN *tḥärgaL* from *əffərgiy* “be afraid,” and AN *tḥaʔgärt* from *uffuʔgər* “be proud (of).” The same applies to some derived verbs, for example, AN *tšäffärt* from *əṣṣəffər* “be next to,” AN *tṣaʔḥ(f)aḍ* (variant *äʔsuʔḥ(f)uḍ*) from *ṣuʔḥ(f)uḍ* “accompany,” AN *tmaṣṣäl~tmaṣṣäyḍ* from *umḡmuyṣiy* “be or become hungry,” and AN *tämrärt* from *ämṣär* “go quickly.”

AN in “*a*—Non-*a*”

Masculine ANs

Some verbs have a single AN in “*a*—non-*a*,” for example, AN *oʔdyuḥ* from *uʔḍuḥ* “admire,” AN *aḍruṣ* from *aḍruṣ* “be accompanied by milk, butter, sauce (dish,

platter); add a condiment,” AN *aṛīh* from *aṛīh* “overcome,” AN *äskər* from *äskər* “make,” AN *ässən* from *ässən* “know,” AN *ässir* from *ässir* “veil,” AN *äsbiy* from *äsbiy* “hinder,” AN *äwTʸuf* from *äTʸif* “let go; lend money,” AN *äwtʸum* from *äTʸəm* “enter,” AN *äʒīg* from *äzīg* “take something and leave quickly with,” and AN *aʾməd* from *amid* “finish.”

In other cases, this is one of the existing variants, often next to a AN in “non-*a*”, for example, AN *aḏḏih* (*/uḏḏih*) from *aḏḏih* “fold,” AN *ažəməḏ* (*/[rarer] užməḏ*) from *ažməḏ* “close,” AN *äššumḏ-äʾšmuḏ* (*/išmuḏ*) from (*ä*)*šmuḏ* “shower,” and AN *äʾšnīy* (*/išnīy*) from *əšnīy* “be smart.”

Feminine ANs

While *ažəDʸ* (variant *užəḏ*) “being sweet to the taste (*ažəḏ*)” is the only AN in “*a*—non-*a*” from a conjugated adjective, this pattern is quite common for FANs from conjugated adjectives and/or verbs that imply “to become. . .”: see AN *toʾzīh* from *oʾZuf-uʾz(z)uf* “(to be) long,” AN *tyäriʾḏ* (variant *tyärät*) from *yärä* “(to be) yellow,” AN *täšumḏ* (variants *t(ä)šmaḏ/ t(ä)šammaḏ*) from *išmaḏ* “become cold,” AN *taṁzīh* (variant *uṁzīh*) from *mazzūg* “(to be) small” and *uṁzīg* “become small,” and AN *täšmimt* (variant *täššəmmih*) from *šəmuṁ* “(be) bitter” and **əšməṁ* “become bitter.”

However, the said AN is attested in other cases, for example, AN *tyäguḏ* from *äykuḏ* “be ashamed,” AN *tāzəL* from *awzīy* “sneak,” AN *tyäguṁt* from *ätʸkəm* “reach (a place. . .),” AN *tāwrih-tōrih* (uncommon variant *ūrih*) from *ūrih* “work,” AN *taymiʾḏ* from *aymi* “put on henna,” AN *taṣzuḏ* from *aṣzuḏ* “crush on the molars,” AN *tayriʾḏ* from *ayri* “read,” and AN *tānəkt* from *änig* “ride a horse; mount.”

Some FANs have a particular use, such as AN *toʾzūyḏ* from *azzig* “be milked,” whereas *äʾzuzg* is the AN of *azzig* “milk” or AN *täydʸəgt-täyDʸəgt* from *ädʸig* “drive; spell,” whereas *iDʸəg-ədʸg* is only the AN from *ädʸig* “spell.” *tmankurt* is the AN from *unkur* “(get up)” and *ṭaṣzīh* the AN from *aṣzī* “(break) up.”

Finally, this pattern is attested for some trilateral or derived verbs such as AN *täyšbəT-täyäšbəT* (variant *oʾkušbih*), from *əkkušbih* “grow”; AN *tgämkih*, from *ägämkih* “spell-out letters”; and AN *tānīZ* (variant *ānīz*) from *ənnīz* “be close, to approach.”

AN in “Non-*a*—*a*”

Masculine ANs

Only three MANs have this pattern: *uḏaš*, which serves as the (suppletive) AN from (*uṣ*)*šumṁih* “sleep”; *əts(s)än*, a possible variant of *täysäkt* as the AN from *ätʸši* “dress”; and AN *užmah* (preferred to *ažmi*) from *ažmug* “sew.”

Feminine ANs

FANs in “non-*a*—*a*” are less rare. Eight of them have no variant; thus AN *tuʾvräh* from *oʾər*, in the sense of “fill”; AN *tiʾšaḏ* from *äʾši* “buy”; AN *tugāh* from *okkīh*

“carry; be pregnant”; AN *tuyah* from *ällih* “search”; AN *t(ə)närt* from *yinər* “orient”; and AN *tūḍaḍ* from *ūḍih* “repudiate (one’s wife).” Two conjugated adjectives have an AN of this type, *tīnāT* “be new (*äynäh*)” and *tušḥaḍ* “be strong (*aṣḥah*),” while “strengthen” (the AN of *uṣḥih*) is pronounced *uṣḥih*, rather than *tušḥaḍ*.

About the same number have a variant. Sometimes there is a difference, like UFAN *təzzād~təzād* and MAN *əẓih* from *äẓih* “remove the skin.” Most often there is no difference in meaning whether the variant is a MAN or FAN: FAN *tuggāh* (variant MAN *uggih*) from *oggih* “end, dry up”; FAN *t(i)yaṭṢ* (variants MAN *uḥtiṣ* and FANs *ṭhattāS/ṭhat(t)šāh*) from *aḥtiṣ* “cutoff”; FAN *tumḍād* (variant MAN *umḍih*) from *umḍih* “be old; age, wear out”; FAN *tiʷwräh* (variant FAN *taʷrt*) from *uʷwur* “become dry”; and FAN *tiʷgräh* (variants FANs *tiʷgərt/tiʷgrih*) from *oʷgər* “steal.” The same applies to the AN of two derived verbs including FAN *tgiʷWärt* (variant MAN *äzgiʷWər*) from *ižgiʷwur* “betray trust.”

Plurisyllabic AN in “Non-a”

Masculine ANs

Plurisyllabic ANs in “non-a” are common for trisyllabic verbs. Examples are AN *iʷgəTʷi* from *iʷgəTʷi* “be innumerable; increase (for noise)”; AN *əftəntər~uftəntər* from *ufiəntər* “swell, swell”; AN *ignəḍi* from *əgnəḍi* “tremble with cold”; AN *ərgəgi* from *ərgəgi* “shiver”; AN *ūžirih* from *ūžirih* “be separated; part”; and AN *iymunḍur* “be beautiful.”

They are also used for nasal-derived verbs such as AN *unuʷgər* from *unuʷgər* “run away in hiding; deny”; AN *umgəgi* from *umgəgi* “collapse (for a well)”; and AN *əməgyih* from *əmməgyih* “lunch.” For verbs derived in “S,” however, they are exceptional. Let us, nonetheless, mention *əṣmuʷḍi* (variant *äʷṣmuʷḍi*), AN from *əṣmuʷḍi* “hide one’s face”, and *ižḥumḥi* AN from *ižḥumḥi* “carry a child on one’s back.”

Feminine ANs

A trisyllabic FAN appears for only two bisyllabic verbs: AN *təžəšših* from *ätʷši* “eat” and AN *təžəšših* from *äšbi* “drink.” *təžəšših* is in competition with *əžəšši* – one of the two MANs in this bisyllabic verb pattern.

Plurisyllabic AN in “a”

Masculine ANs

This pattern is the only one attested for nine bisyllabic verbs: *əḍrəg* “be polygamous” (AN *ädäräg*), *ayni* “be rich” (AN *äyyanä*), *ämšuḍ* “comb” (AN *ämäšaḍ*), *ässəf* “pluck” (AN *ässäḍäf*), *aḍyuṣ* “walk by night” (AN *ädayaṣ~aṣaDʷaḍ*), *äkfīy* “climb” (AN *ägäffäy*), *äktub* “write” (AN *ägättäb*), *ärtiy* “cover (a dead person)” (AN *ärättäy*), and *äzzīy* “swear” (AN *äzäddäy*). But this trilateral pattern frequently alternates with a biliteral pattern:

- A FAN in “a” for *ädgən* “grease” (ANs *ädäggän/tdägnäh*), *ägwiḥ* “bellow” (ANs *ägäbbäh/tgäwät*), and *äyṁuḗ* “be thin” (ANs *äyammaḗ/tyämmaḗ*).
- A MAN in “non-a” for *ädṁər* “speak” (ANs *ädämmär/ädṁər*), *äfrəd* “pick” (ANs *äffäräd/ufṛəd*), *ögri(i)ḥ* “hear” (ANs *ägäräh/ugriḥ*), *ägyub* “be ravenous” (ANs *ägäyäh/ugyub*) – as AN of *ägyub* ‘having rabies’, *ugyub* alternates with *tgäybäh*, *äyṁiḥ* “be angry; anger” (ANs *ayannäh/uyṁiḥ*), *ayyiḗ* “forget” (ANs *ayalläd~ayayäd/[rarer] uyyiḗ*), *änšuḗ* “fall” (ANs *anaṣṣaḗ/unšuḗ*), *ärṁən* “make dirty” (ANs *äräffän/ärṁən*), *ädṁin* “make commitments” (ANs *aḏammän/uḏṁun*), *äsṛəm* “level” (ANs *ässäräm/äsṛəm*), and *ällig* “be calm” (ANs *ättäyäg/öllög*) – with the meaning “become calm; stand still”, *ägyub* has for AN *ädäyäg* or *älläg*.

Only a few ANs in “a” correspond to trisyllabic or derived verbs: AN *äybaḏa* (irregular) from (*iṣ*)*šiybuḗ* “entrust”; AN *äyaD’äḗ* from *ayaD’äḗ* “show laziness”; AN *äddäwäm* from *ädämäh* “go on; be of permanence”; AN *ä’ḗḗa’war~ä’ḗḗa’war* from *əḗḗi’wur~əḗḗi’wur* “render thick; ennoble”; and AN *a’gaḏah* (variant *ä’guḏiḥ*) from *əgguḏiḥ* “be or become drowsy.”

Feminine ANs

täffa’gägt is the AN of the adjective *fö’gug* and the verb *äffa’gäg* “have buck teeth.” However, apart from AN *tärka’gämt* (variant *ərki’gim*) and from *urku’gum* “be calm” and AN *ta’maḡqaS* from the Arabicized verb *a’maḡqaš* “struggle before death,” the few plurisyllabic FANs in “a” correspond to bisyllabic verbs, such as AN *tänändräh* (~*tnändräh*) from *ändər* “burn,” AN *tayayäS* from *ayyiš* “be safe and sound,” and AN *täzägärt* (one of four) from *äzgər* “sprout”: note AN *täsäkräh* from *äskər* “relieve oneself,” while it is AN *äskər* for the less specific sense of “do.”

AN in “a—a—Non-a”

Masculine ANs

The ANs in the “a—a—non-a” pattern are those of 17 bisyllabic verbs. This pattern is the only one attested for six verbs: *änwi* “cook; ripen” AN *änäbbi*; *aṛḏiy* “lend” AN *aṛaḏiy* (*urḏiy* means “made to measure,” but this meaning has not been noted for the verb *aṛḏiy*.); *ärkiy* “be defiled; defile” AN *aṛaguy*; *äymi* “be softened; soften” AN *äyämmi*; and *aḗri* “find” AN *aḗari*; *əḗriḥ* “hunger (animal)” AN *äḗariḥ*.

These ANs can alternate with other MAN patterns (most often in “non-a”): *aṭṣi* “laugh” AN *aḏaṣṣi*/[less good] *uṭsi*, *änki* “smell” AN *änäkki/änki*, *änti* “sting” AN *änätti/änti*, *azḗig* “strangle” AN *aḗäD’ig*/[less good] *uzḗig*, *ayriš* “be sure” AN *ättayrəš/uyriš*, *aṛiḥ* “increase” AN *ässṣaṛiḥ/aṛiḥ*, *ärwiḥ* “be stirred up, and stir” AN *äräwiḥ/äräbbäh*.

In a few cases, this pattern has a FAN variant. This feminine provides a UAN for three verbs: *aṛḗi* “break” MAN *aṛaḗi* and UFAN *ṭraḗiḥ*; *aṛḗum* “untie; pay” MAN *aṛaḗəm* and UFAN *täṛaḗmih~t(a)ṛaḗəmt*; and *ätfi* “spill” MAN *ädäffī* and UFAN

tädäffīʔd. It is a simple variant for *zzuʔr* “see” MAN *ažari*/FAN *tmaʔdärt* ~ *tämmiʔdräh*. As for *äzri(i)h* “follow,” it has two MANs *äzäri/əzri* and one FAN *tžärät*.

Only five plurisyllabic verbs have a MAN of this pattern: *ruʔri* “vomit” AN *äraʔri*, *ižəʔzi-ižəʔzi* “shout” AN *aʔžəʔzi-aʔžəwž*, *ämṣuynāh* “get angry” AN *äymaynīh/ äʔṣuynīh*, *əššəgi* “miss a day” AN *äʔš(š)ägi*/[better] *äʔš(š)əgi*, and *ədduriy* “become one-eyed” MAN *äʔdürwi* and UFAN *tädyäl*.

Feminine ANs

Among the dozen or so bisyllabic verbs with this FAN, there is *ažum* “untie; pay” and *äffi* “reverse,” which have (cf. previous section “Masculine ANs”) an UAN in “*a—a—non-a*.” In four cases, the FAN is attested alongside a MAN in “*non-a*”: *əđbi* “do penance” FAN *täʔdəbiʔd* (*uđbi~əđbi*), *äbđih* “walk” FAN *taʔwäđih* (*ibđih/ ižigž~ižigž*), *äzgi* “take” FAN *täzöggiʔd-täzəwggīʔd* (*izgi*), and *äddih* “go astray” FAN *taʔwäđih* (It is *əđbi* in the sense of “die” which has for AN *uđbi/əđbi*).

In six cases, the verb has only one FAN (or two): *äžar* “be impregnated (female)” FAN *täžärīh*, *änši* “extinguish” FAN *tänässiʔd* (*änši* “spend the night” has for AN *tänäskt*), *änzi* “forgive” FAN *tänäzziʔd*, *ättar* “ask” FAN *taʔwäđrih-täwäđrih*, *äžyih* “throw away” FAN *täžäyət-tžäyät*, and *äykuđ* “be ashamed of” FAN *täyäguD-tyäguD*.

Only two trisyllabic verbs have an AN in “*a—a—non-a*”: *aššalli* “pray” FAN *taʔžalliʔd* and *əkkušbih* “be great; grow” FAN *täyäšbət-täyšbət* (*oʔkušbih*).

AN in “*a—Non-a—Non-a*[—Non-a]”

Masculine ANs

Two cases can be distinguished, depending on the presence, or absence, of a laryngeal – after the initial *ä*-.

- (a) The “*a—non-a—non-a*[—non-a]” pattern is very frequent with a laryngeal.

For trisyllabic verbs and nasal derivatives, this pattern may be either a variant of another pattern, usually the one in “*non-a*,” or the only pattern used. Examples are AN *äʔṣugrih* (*uṣṣugrih*) from *uṣṣugrih* “return,” AN *äʔṣuđriš* (*uṣuđriš*) from *uṣṣuđriš* “die without having one’s throat slit,” AN *äʔnuʔṣəš* (*ənuʔṣəš*) from *ənnuʔṣiš* “love,” AN *äʔmənšəh* (*əməšəh*) from *əmmənšəh* “dine out,” AN *äʔburɣiy* from *uḃburɣiy* “give zrig or very diluted tea,” AN *oʔburbuđ* from *ubburbuđ* “become spotted,” AN *äʔmuktər* from *uʔmuktər* “ride on the rump,” AN *äʔfuɣyih* from *uffuɣyih* “be bald,” AN *oʔfušši* from *affašši* “be silent,” AN *oʔkušbih* from *əkkušbih* “become tall,” AN *äʔnərəm* from *ənnərəm* “writhe in pain,” AN *äʔgüđer* from *uggüđer* “invite oneself to eat,” AN *äʔgəngiyi* from *əggungiy* “itch,” AN *äʔkiʔgiš~äʔkugīgiš* from *ikkīgəš* “become blind,” AN *äʔkugīh* from *əkkuḡih* “burn (food),” and AN *äʔkənfih* from *əkkənfih-ukkunfih* “take a rest.”

For “*S*” derivatives, the *äʔ*- pattern is practically the rule, for example, AN *äʔšfuɣyih* from *əšfayyāh* “make bald,” AN *äʔškiʔgiš* from *äškaʔgäs* “make blind,” AN *äʔžguđih* from *əžguđih* “put to sleep,” AN *äʔšquffih* from *əšruʔyiy* “anger,”

AN *äʕswuḏīh* from *äswuḏīh* “send,” AN *äʕsnūḏīh* from *äsnūḏāh* “separate,” and AN *äʕškunḏīh* from *äškənḏīh~äškunḏīh* “force to take a break.”

Note that this *äʕ*-prefixed pattern is sometimes used as an AN in *m* of bisyllabic verbs, such as AN *äʕmərḏīh* from *urḏīh* “be happy” – and AN *äʕmunḏug* (*unḏug*) from *anḏug* “taste.”

- (b) The “*a*—non-*a*—non-*a*[—non-*a*]” pattern is very rare without a laryngeal. It is found for some bisyllabic verbs but never as the only variant: AN *ättükiy* (*ükiy*) from *ükiy* “love,” AN *äzəggər* (*əzəgər*) from *əzəgər* “go out,” and AN *äššəlli(i)h* (*[better] əštʕig*) from *äštʕig* “curdle (for milk).” For trisyllabic verbs, on the other hand, it is the only AN used in two out of three cases: AN *äffiʕri* from *əffiʕri* “fall apart (for a hole),” AN *äzəzrih* from *əzəzrih* “make follow,” and AN *äzgiʕWər* (variant *tgiʕWärt*) from *izgiʕwur* “betray trust.”

Feminine ANs

This pattern is attested for some bisyllabic verbs but without the laryngeal in the first syllable (note that three times out of four there is one in the second syllable): AN *taḏuʕrih* (*uḏur*) from *aḏur* “fall; raid,” AN *täguʕrih* from *ägur* “draw water from a well; rise (star),” – *täguʕrih* has for variant *təgrih* in the first sense and for variant *toʕgurt* in the second sense – AN *tämuʕri* from *umuʕr* “get older,” and AN *täš(š)əddih* from *əštīd* “become slim” and UAN *tändiywih* (AN *unḏuy*) from *ändiy* “entomb.”

Besides *tässədbəḏ*, the AN of the “*S*”-derived verb (*əs*)*sədbih* “leave in the afternoon,” two other ANs correspond to both a trisyllabic verb and an adjective: AN *täššəmmih* (*täšmim*) from (*ət*)*təšmim* “become bitter (*šəmmum*)” and AN *täzḏibbiʕḏ* from *izḏibbi* “turn red (*zəbḏā*).”

Other Cases

Monosyllabic in “Non-*a*”

Eight MANs are in “non-*a*,” but only one corresponds to a monosyllabic verb: MAN *ūr* (*tārt*) from *ūr* “associate.” Apart from *grər* which serves as AN for *əssigrəḏ* “drive,” all the others correspond to bisyllabic verbs: MAN *əbg* (*uḥḥug*) from *uḥḥug* “be far away; move away,” MAN *ədʕg* (*/iDʕəg/täydaʕgt*) from *ädʕig* “spell,” MAN *ḥkun* from *ufkun* “grow old,” MAN *məḏ* (*/aʕməḏ*) from *amiḏ* “finish,” MAN *shur* from (*ə*)*shur* “absorb something at daybreak (when fasting),” and MAN *üktʕ* (*/ükiy/ättükiy*) from *ükiy* “love.”

Only two FANs are in “non-*a*”: FAN *tīgt~tīkt* from *tīg* “hurt; make hurt” and FAN *tītṣ* (*/t(i)yaṭṣ/thattäS/uḥtiṣ*) from *aḥtiṣ* “cut.”

Monosyllabic AN in “*a*”

The only MAN in “*a*” is *ṛkaʕz*, AN of the trisyllabic verb (P) *yarkaʕza* “grill (meat).”

Nine FANs are in “*a*.” Four of them correspond to monosyllabic verbs and five to bisyllabic verbs: on the one hand, FAN *taʕD* (*/taʕädt/tuʕuḏ*) from *aʕḏ* “tend,” FAN *tārt* (variant *ūr*) from *ūr* “join,” FAN *toʕfi* from *uʕf* “swell,” FAN *taʕfi* from *aʕf* “yawn,”

Borrowings from Arabic

Three FANs in “*a—a*” are clearly borrowings from Arabic: FAN *käyyä* from *äkwäh* “cauterize,” FAN *raqqa* from *äryah* “blather,” and FAN *šarṭa* from *äšraṭṭah* “scarify.”

There are also a number of borrowings from Arabic, but these do not constitute nonintegrated borrowings.

Summary Table of the Main Patterns

The frequency (Fq) of each of the vowel patterns is given according to the gender of the AN, but it is relative and takes into account the fact that FANs are much less numerous than MANs. The frequency is then specified in each column according to the number of syllables of the verb. Here again, it must be considered that bisyllabic verbs (Bi) are more numerous than trisyllabic or derived verbs (Tri), and monosyllabic verbs (Mo) are very few in number.

AN	M/F	Frequency	Mo	Bi	Tri
'non- <i>a</i> —non- <i>a</i> '	M	Fq+++	—	+++	2~3
	F	Fq+	3	27	—
' <i>a</i> — <i>a</i> '	M	Fq	1	17	1
	F	Fq++	1	++	9 T+
' <i>a</i> —non- <i>a</i> '	M	Fq+	—	26	—
	F	Fq+	—	23	3
'non- <i>a</i> — <i>a</i> '	M	very rare	—	3	—
	F	Fq	—	16	2
'non- <i>a</i> —non- <i>a</i> —non- <i>a</i> '	M	Fq++	—	—	++
	F	very rare	—	3	—
' <i>a</i> — <i>a</i> — <i>a</i> '	M	Fq+	—	29	5
	F	rare	—	5	3
' <i>a</i> — <i>a</i> —non- <i>a</i> '	M	Fq	—	17	4
	F	Fq	—	12	2
' <i>a</i> —non- <i>a</i> —non- <i>a</i> [—non- <i>a</i>]'	M	Fq+++	—	4	+++
	F	rare	—	5	3

The only relatively frequent patterns in all cases are the “non-*a*” patterns: “non-*a*—non-*a*” for MAN/FAN from bisyllabic verbs and “non-*a*—non-*a*—non-*a*” for MAN from trisyllabic and derivative verbs.

With the exception of rather average frequency patterns such as “*a*—non-*a*” and “*a*—*a*—non-*a*” (which are attested almost equally for both genders but almost exclusively as ANs from bisyllabic verbs), all other frequent patterns (Fq++ and Fq+++), are, only so, compared to a particular gender (“*a*—*a*” for FANs and “*a*—*a*—*a*” for MANs from bisyllabic verbs) or compared to a particular type of verb (“*a*—non-*a*—non-*a*[—non-*a*]” for trisyllabic or derived verbs and especially for “S” derivatives).

Analysis and Comments

I will come back to the prevalence of certain patterns in section “[From a Morphological Point of View](#)” and will end with some syntactic observations. Before that, I will look at the variations in some ANs from the point of view of radical consonants.

From the Point of View of the Root

ANs tend to exhibit significant irregularities and some aspects are discussed here.

Degemination

The AN often has the same geminate root as the corresponding bisyllabic verb, e.g., *uggih* “need,” AN *uggi*; *äggur* “pull” and *uggur* “be barren,” AN *uggur*. But this is not always true. In about 15 cases, the AN (or one of its variants) has a simple radical and not a geminate, as in the imperative: for example, *uḥḥug* “be far away. . .,” ANs *əbg/uhḥug*; *ädʷig* “spell,” MANs *ədʷg/iDʷag*—variants [rarer] of FANs *täydʷagt/täyDʷagt*.

Correspondences such as *ḍ/dd*, *y/qq*, and *y/ll* are consistent with Zenaga phonology: *äddəm* “set (star); fall (drop),” AN *uḍum/əddəm*; *äddəž* “grind,” AN *uḍəž*; *adḍuḍ* “be suckled; suckle,” AN *uḍuḍ*; *aḍḍuf* “possess,” AN *uḍuf*; *aqqiy* “look at” AN *uyiy~uyuy*; *älli* “lick up,” AN *uyi/əlli*; *ällih* “seek, to search,” [irregular] FAN *tuyah* – the geminate is frequently found in other Berber dialects but with exceptions such as *uḍum* “ooze” in Rif and *idim~udum* in Kabyle (Nait-Zerrad 1998–2002: II, 336–7); we also, sometimes, find the same alternations between the verb and the AN (see under *ḌF*, Nait-Zerrad *id.*: III, 447–8).

The *ž/zz* correspondence is also regular in *zzuʔr* “see,” AN *ažari* (>*ažari*?-n. . .), but it does not directly explain the shift in position of the laryngeal.

The *g/kk* alternation is not regular in Zenaga, but it regularly accompanies degemination, cf. MAN *ugi* from *äkki* “pass (through)”; MAN *uguḍ* from *äkkuḍ* “be smooth; lose one’s hair; remove hair”; MAN *ugum* from *äkkum* “be sunken; drive (a stake); strike”; FAN *tugāh* from *okkih* “carry; be pregnant”; and FAN *tugāh* from *oggih* “end up” (MAN *uggih*).

To all these relatively simple cases, we should also add those which correspond to de-assimilation as in FAN *taʷwādih* from *äddih* “go astray” and FAN *taʷwādr̥ih~täwādr̥ih* from *ättar* “ask,” suggesting, on the one hand, *w + d > dd*, and on the other, *w + t > tt* (with AN forms less “evolved” than the verbal forms).

Gemination

Gemination, on the other hand, which is absent from the “on-augmented” forms of the verb (I, A, P, and NP), is sometimes present in the AN. This is notably the case of MANs in “*a—a*” such as MAN *äggäy* (variant *ugiy*) from *ogiy* “hang” and MAN *äbbäš* (variant *uʷwəš*) from *äwuš* “help”; in “*a—a—a*” such as MAN *äzäddäy* from *äzziy* “curse” or MAN *ässäräm* (variant *əsrəm*) from *äsrəm* “clear, level”; and in

“a—a—i” such as MAN *ädäffī* from *ätfi* “topple over.” This is also the case for some FANs in “a—a” like FAN *tgämmärt* from *ägṃur* “be narrow” (MAN *ugṃur*).

Other cases of gemination appear more isolated, and these are, for example, for MANs: MAN *äzəggər* (variant *əzgər*) from *əzgər* “go out,” MAN *adḏih* (variant *uḏih*) from *yaḏih* “fold,” and MAN *əṣṣagaḏ* (variant *uṣkuḏ*) from *aṣkuḏ* “pound; prune.” This is the case, also, for some FANs: FAN *tumṃuṣ* (without *ʔ*) from *aʔmuṣ* “bite”; UFAN *tuʔṃmuḏ* from *uʔmuḏ* “close (of itself)” and FAN *tāwṃmuḌ* from *āmīy* “move toward” (more frequent MAN *umuy*). Also see FANs with geminate *m* or *n* (such as AN *təmmiḏah* from *əmiḏ* “long for”) in previous section “ANs in ‘Non-a—Non-a[—Non-a]—a’”. With the exception of the last two cases, the others are explained either by the tendency to unvoice in contact with a voiceless consonant, or by the tendency to voice between vowels.

I will return to some cases of gemination in section “IA Patterns”.

Velars

Sometimes, a voiceless velar (*k*) in common Berber corresponds to a voiced velar (*g*) in Zenaga in the various lexemes of the root, for example, in MAN *ugruf* from *ägruf* “retract” for GRF (Berb. KRF) or MAN *ugyi* from *ogyi* “pass the meridian” for GY? (Berb. KL).

However, there is often an alternation between *g* and *k* (or even *g*, *k*, and *w*, as in FANs *tärwäh/tärukt* from *ärug* “give birth”). The examples below can be added to the preceding examples of degemination: FAN *tänəkt* from *änig* “ride a horse,” MAN *ägäffäy* from *äkfiy* “go up,” MAN *ägättāb* from *äktub* “write” (< ar. KTB), and ANs *ugṃuk* and *tgämkäh* from *ägṃug* “follow” (berb. GMK).

At times, the velar *k/g* disappears in ANs. In final position, it seems to change to *h* in MAN *umṣih* from *umṣig* “become small” and in MAN *uṣmah* from *aṣṃug* “sew” (FAN *təṣṃmāgḏ*) – note that we find both *k* (Ghadamsi *ezmæk*, Lanfry 1973: n° 1806) and *y* (Tuareg *əzməy*, Prasse et al. 2003: 921) in other Berber dialects; for the alternance *y/k* in the last radical, see Kossmann (1999: 186–192) – and MAN *ägruḥ* (*/ugrug*) from *ägrug* “preserve; pour through an orifice.” The change of the first radical in *y* (cf. Kossmann 1999: 192–3) is attested in the following: MAN *äynāš* (*/ugnāš*) from *ognāš* “argue” (Berb. KNS), MAN *äyṛaḏ* (*/ugruḏ*) from *ägruḏ* “scrape (to clean),” FAN *täyṣbət~täyṣbət* from *əkkuṣbih* “grow,” and FAN *täyṣäkt* from *okši* “graze.”

The *k* in *täyṣäkt*, more surprisingly, is found in the FANs of three other verbs with a final laryngeal root: FAN *täyṣäkt* from *ätʔši* “dress,” FAN *tänsäkt* from *änši* “spend the night,” and FAN *taʔškäkt* from *aʔzi* “build.” This *k* is explained, for the last root, by a *k* (*/q*) attested in some Berber languages. This is not the case for the other roots; however, the *k* may be related to the final *-u* of ^W*äksu* “graze” and ^{WY}*äṃsu* “spend the night” in Tuareg (Prasse et al. 2003). There is no *-u* in *ls* “dressing” in Tuareg, but in Zenaga, there could be contamination between the ANs “grazing” and “dressing,” as they have become homonyms.

Glottal Stop

As the final radical, the glottal stop has a few particularities: absent in absolute final position, it normally reappears before a suffix, for example., (ŠG?) *äʔš(š)əgi* “be absent; absence” but *äʔš(š)əgiʔ-n-k* “your absence”; (NTB?) *əntəbi* “calmness,

tranquility” but *əntabiʔ-n-s* “its calmness” (For FANs, see below). All MANs end in *-i* except two irregular ANs which end in *-a*: *aṣṭa* “braiding of hair; weaving (of a mat)” (ZZʔ, Berb. ZD/ZD) and *äyṣanä* (*äyṣanäʔ-n-s* “his wealth”), AN from *ayni* “be rich; grow rich” (YŇʔ < Ar).

In contrast, the presence of the glottal stop is less regular word initially, sometimes appearing in the AN while absent in the verb. In MAN *iʔriš* (*/iriš*) from *ärəš* “be torn; tear (from itself)” (Berb. YRS), it represents a missing radical. The same is probably true of MAN *uʔwəš* (*/äbbäš*) from *äwuš* “help” and MAN *aʔməd* from *amid* “finish.” There may be assimilation with derivatives, especially in “S,” with prefix *aʔ-*, as in MAN *äʔšmuḍ* (*/äššumḍ/išmuḍ*) from (*ä*)*šmuḍ* “shower” or MAN *äʔzīg* from *äzīg* “pick up something and quickly leave with.” Finally, the presence of the glottal stop may remain unexplained as in MAN *aʔyāš* from *ayūs* “clean” (<Ar. YBSʔ).

Other Cases

Some rare forms are more or less integrated borrowings from Arabic, such as *äl-faraḍ* > *ätʔfaraḍ* (*/afraḍ/ufṛuḍ*) “obligation” (< Ar. FRD), *äl-dāwām* > *äddāwām* “continue, be of permanence” (< Ar. *ḍāmma*), and *ätturyāš* (*/uryuš*) “be cheap” (< Ar. RḤṢ).

Some irregularities pertain to more or less pan-Berber substitution phenomena, such as MAN *innih* (same root as P *yənnāh*) from *əzzən* “say,” FAN *tmättänt* (theme partly similar to that of IA *yətmättāh*) from *ämṃih* “die,” MAN *ižīgž-iʔžigž* (variant *ibḍih*) from *äbḍih* “walk; go” – cf. *əžəgəž* “trip” in Tətsərret (Attayoub 2001: 97) and MAN *iʔḍi* from *äwih* “bring in” (WH/H*, Berb. WY), and *äwih* “strike” (WH/H*, Berb. WT).

The presence of an *m* is reminiscent of a borrowing from a nasal derivation theme. From this point of view, the MANs *äʔmərḍih* (from *urḍih* “be happy”) and *äʔmunḍug* (from *anḍug* “taste”) are regularly formed as is the UFAN *tmankurt* (from *änkür* “stand up”), while the FANs *tmaʔḍärt-tämmiʔḍräh* (from *zzuʔr* “see”), *tmäntänt* (from *aʔni* “kill”), and *täniykāḌ* (from *äykih* “despise”) are not.

In some cases, the evolution can be explained quite simply: emphatization of *ss* (geminate counterpart of *š*) by *ḍ* for MAN *anaṣṣaḍ* (*/unšuḍ*) from *änšuḍ* “fall”; double derivation of TLG (>LLG), with or without gemination of the first radical T (in Zenaga, *ḍ* is not only the non-geminate correspondent of *dd* but, also, that of *tt*, especially when */t/* is intervocalic), for *ättäyäg* (*/əlläg*) “be calm” and *ädäyäg* (*/əlläg*) “become stable.”

In other cases, the evolution is more complex and pertains to Kossmann’s (1999: 120–125) observations on the realizations of **Ĥ* (**ḥ*). In Zenaga, a long vowel can alternate, either with *f*, as in UFAN *tikiḥ* from *äkfi* “give” (MAN *ukfi*), or with *w*, as in MANs *äddäž/īddiž* from *äwdəž* “lie down,” MAN *äggäš* (*/ügəš*) from *äwgəš* “gird,” and MAN *īddih* (*/tīddih*) from *äwdəḍ* “stand up; halt” (cf. Taine-Cheikh 2015: 303–304).

Many ANs seem irregular in Zenaga, and comparison with other languages is often enlightening. To the previous cases, we can add that of *tugguḌ*: although the *z* of the FAN is absent in the verb *oʔgəž* “know,” it appears elsewhere (cf. Tuareg ^{WY}*agəz* “protect,” Prasse et al. 2003: 260).

From a Morphological Point of View

AN patterns will first be considered in their relationship with similar verbal patterns. The study will then focus on FAN endings.

Negative Perfective (NP) Patterns

Among the most frequent AN patterns, there are those of ANs in “non-*a*,” and, foremost, those of bisyllabic verbs. Unlike FANs, MANs in “non-*a*—non-*a*” are numerous and share the same pattern as the NP of verbs. For the vast majority of bisyllabic verbs, only the NP is in “non-*a*,” as opposed to the “*a*—non-*a*” pattern of I = A, and the (inverted) “non-*a*—*a*” pattern of P. For internal verbs, this pattern is also that of I = A (see MAN *uḥḥug* from *uḥḥug* “be away; go away”).

For trisyllabic and derived verbs, there are almost only ANs with “non-*a*” thematic vowels, but they fall into two groups: either the pattern includes only “non-*a*’s, or the prefix vowel is replaced by *a*²-. When it includes only “non-*a*’s (prefix + thematic vowels), it regularly merges with the common NP = I = A pattern, for example, MAN *ərgəgi* from “shiver” I *ərgəgi* A = NP *yərgəgi* (as opposed to P *yərgägä* and IA *yəttərgəgi*) and MAN *əmdukkīy* from “become friends” I *əmdukkīy* A = NP *yəmdukkīy* (as opposed to P *yəmdukkāy* and IA *yəttəmdukkīy*). These include the MANs of “*S*” derivatives beginning with *a*²-, for example, MAN *ä’sədrər* from “cut into strips” I *sədrər* NP = A = IA *yəssədrər* (as opposed to P *yəssədrär*) and MAN *ä’fuyyih* from “be bald” NP = A *yuffuyyih* (as opposed to P *yäffayyäh* and IA *yətfuyyih*).

These “non-*a*—non-*a*” thematic vowels are common to the ANs of nasal derivatives and to the most frequent pattern of habitual nouns in *m/n* (Taine-Cheikh 2019: 191–200) such as *ämu’gər* “thief” and *änfərgiy* “fearful, always afraid.”

The glottal stop present in *a*²- tends to function in Zenaga as a specific (M)AN marker for trilateral and derived verbs, as if the subspecification of the NP-scheme made it necessary to add a distinguishing marker.

IA Patterns

The patterns of the IA seem to have influenced those of some ANs. The IA of bisyllabic verbs, among others, is mostly formed with the prefix *t(t)*-. However, it is also formed by gemination (doubly) of a radical (IA-D), and the two processes can coexist: *aḡniš* “put on fibulae,” IA-t *yəttaynāš*/IA-D [more frequent] *yäḡannāš*; *äyāš* “mow,” and IA-t *yəttäyāš*/IA-D *yällāš*. Sometimes, there is also no gemination, but only addition of a vowel after the first radical (IA-v): *ədrəg* “be polygamous,” IA-t *yittədrəg*/IA-v *yidäräg*, and *ogri(i)h* “hear,” A-v *yəgäräh*.

If *ättükīy*, the AN from *ükīy* “love well” and *əttūih*, the AN from *ū’ih* “be awake; awaken,” owe their form to those of IA-t *yəttükīy* and *yəttū’ih*, these can only be exceptions (perhaps due, at least in the second case, to Arabic influence). However, there are many cases of convergence between ANs in “*a*—*a*[—*a*]” and IAs in “*a*” (or, more precisely, in “*a*—*a*[—*a*]” like *yäḡannāš* and *yällāš*, or in “non-*a*—*a*—*a*” like *yidäräg*). These convergences appear in the respective ANs of the examples given above: *aynāš* (*uḡniš*), *äyāš*, *ädäräg*, and *ägäräh* (*uḡrīh*). They are also found

in MAN *ānaẓ* from “remove liquid” (IA-D *yivānnaẓ*), MAN *āmāšaḍ* (/imšuḍ/) from “comb” (IA-D *yimāššaḍ*), MAN *aḫraḍ* (/uḫruḍ/) from “be obliged; oblige” [< Ar.] (IA-D *yiffaraḍ*/IA-t *yəttuḫruḍ*), MAN *aṣṭa* “braid hair” (IA-D *yīṣṣaṭṭa*), MAN *āynās* (/ugnās/) “argue” (IA-D *yikkānnās*), and MAN *āyraḍ* (/ugruḍ/) “scrape (to clean)” (IA-v *yigaraḍ*). They are still found in the following cases where a *y* appears rather irregularly in IA and/or AN: MAN *āybbāḍ* from “jump” (IA-y *yāybbāḍ*) and MAN *āyyanā* from “be rich” (IA-D *yīyannā*).

Let us now look at the cases of the FANs. There is no convergence with the IAs for feminines in “*a—ā*” which regularly provide, as I have shown, (F)ANs from conjugated adjectives and, possibly, from corresponding quality verbs, such as *ṭharḫūft*, FAN from *ḫarḫuf* “(be) hard” and *uḫḫurḫuf* “become hard.” However, other FANs in “*a—ā*” (often variants of MANs in “non-*a*” not repeated here) can be correlated with an IA other than IA-t.

Some FANs in “*a—ā*” are correlated with IA-Ds: UFAN *tẓamḍah* from “close,” FAN *tẓamḡaḍ* from “sew,” FAN *tmāntānt* from “kill,” FAN *tdāgrāh* from “be decorated; decorate,” FAN *tdāmnāh* from “let flour rest with yeast,” FAN *ṭayyāT* from “wait,” FAN *tāysākt* from “graze” (IA-D *yikāssā*) and “dress” (IA-D *yiyāssā*), FAN *tgamẓāh* from “scrape,” FAN *tgānāh* from “barrack,” FAN *tkāššāft* from “shame,” FAN *tkāsrāh* from “go down,” FAN *tānsākt* from “spend the night,” FAN *tnāẓāḍ* from “be sold,” FAN *ṭaṣṣaD* from “stink,” FAN *tāzāgārt* (/tāzēgrāh/) from “go outside,” FAN *tāsākrāh* from “relieve oneself,” FAN *ṭayayāS* from “be alive,” and FAN *tẓāyāT* (/tāzēyāT/) from “throw away.”

Two NAFs in “*a—ā*” are correlated with IA-vs [verbs < Ar.]: FAN *ṭayyāT* (/ṭayyiḍ/) from “become expensive and difficult” (IA-v *yīyayā*); FAN *ṭayayāb* from “become dizzy” (IA-v *yīyayāb*).

Some FANs in “*a—ā*” are correlated with IAs with rare prefixation in *y*- (IA-y) or *n*- (IA-n): FAN *tgāmkāh* from “follow” (IA-y *yāygmāg*), FAN *trāwyā* from “run away” (IA-y *yāyrywāy*), FAN *tānāndrāh*/tnāndrāh from “burn” (IA-n *yināndār*), and FAN *tnāndrah* from “steal” (IA-n *yinanḍar*/IA-y *yāyndar*).

By their exceptional character, the last two examples support the correlation between the form of the IA and that of the AN. However, there are many counterexamples: on the one hand, about 20 ANs in “*a*” whose IA is in *tt*- and, on the other hand, a much larger number of verbs, which do not form their IA in *tt*- but, also, do not have an AN in “*a*.”

The second set of counterexamples may be partly explained by the decline of a less productive pattern, especially for MANs. But there is also the fact that (with two exceptions: *aṣṭa* and *āyyanā*) the final vowel of MANs is always *-i* in the case of the last laryngeal radical, which explains the replacement of “*a—ā—ā*” by “*a—ā—non-a*” in MAN *ādāḫfi* from “spill,” MAN *aḍaṣṣi* from “laugh,” MAN *ānākki* from “smell,” MAN *ānātti* from “sting,” MAN *ānābbi* from “cook; ripen,” MAN *aṣaḫi* from “break,” MAN *āyāmmi* from “be softened; soften,” and MAN *aḫari* from “find.”

All verbs, apart from the last (“find,” IA-v *yāḫara*), form their IA by geminating (IA-D). This also explains, besides UFAN *tādāḫfi*ḍ vs. MAN *ādāḫfi*, some feminines like FAN *tānāssi*ḍ from “extinguish” (IA-D *yināssā*), FAN *tānāzzi*ḍ from “forgive”

(IA-D *yinäzzä*), FAN *täʔdäbiʔd* from *ədḥi* “do penance” (IA in *tt-*), and FAN *taʔzalliʔd* from the trisyllabic verb *aṣṣalli* “pray.”

There are very few “*a*—non-*a*” ANs for roots with a final glottal stop (let us, nonetheless, quote FAN *taʔmiʔd* from *aʔmi* “be dyed with henna; put henna,” IA-D *yīyammä*; and FAN *taʔriʔd* from *aʔri* “read,” IA-v *yīyarä*). In roots with final radical H/T, H*, or Y, the AN is quite often in “*a*—[*a*]non-*a*” but without the regularity observed for ʔ.

Identical Imperative and Aorist Patterns

The “*a*—non-*a*” pattern, which is the most common I = A scheme for bisyllabic verbs, is also an AN scheme (see section “AN in “*a*—Non-*a*””). However, rare are the cases where the ANs merge with I = A, even among MANs.

FAN Endings

In principle, feminine nouns are characterized, in the singular, by the addition of two *t*-affixes, a prefix, and a suffix (i.e., a circumfix). While the dental prefix is regularly realized [t], the realization of the suffix is more complex as it depends on the final radical consonant (Taine-Cheikh 2002: 430–434). It is added after the radical consonants *r*, *f*, *m*, *n*, *k*, and ʔ. It is realized *t*~*d* after *g*, *y*, and *T* or *d* in lexemes ending in *h* (there has been some confusion between *T and *H, on the one hand, and *H and *H*, on the other, I leave this problem aside here). Finally, in other cases, it assimilates to the final radical consonant, giving rise to a geminate realized as a tensed consonant (noted by a capital letter), thus *d* + *t* > *D*, *d* + *t* > *D*, *y* + *t* > *L*, *š* + *t* > *S*, *ž/ž* + *t* > *Z*, and *z* + *t* > *Z*.

Among the FANs, a number of them follow these formation rules. This is especially the case for FANs from (conjugated) adjectives in “*a*—*a*” such as *tbäyḏayḏ*, *tbarbaD*, *tkäbbänt*, etc. (see section “Feminine ANs” under “Bisyllabic ANs in “*a*””). This is also the case with many other FANs, whether they correspond to (conjugated) adjectives, such as *tinäT* “being new (*äynäh*),” and *tušḥad* “being strong (*aṣḥah*)”, or not, such as FAN *tukkuṃt* “running,” FAN *təddəS* from “prick,” FAN *taʔzuZ* from “crunch,” FAN *tūḏad* from “repudiate (one’s wife),” UFAN *tgäwät* from “bellow,” FAN *tyayät* from “wait,” and FAN *tuzziʔd* from “be braided; plait (hair).”

Among the FANs without suffix *-t* (assimilated or not), many of them present a single vowel (usually short: /a/ or /i/, never /u/), followed by a more or less reduced *h* which is deleted when annexed. These endings are likely to occur with all radicals.

- (i) Examples of FANs in *-ah/-äh* are FAN *tzamḏah* from “close,” FAN *tgämkäh* from “follow,” FAN *tḏägräh* from “be decorated; decorate,” FAN *tḏämnäh* from “let flour rest with yeast,” FAN *tgämkäh* from “follow,” UFAN *tgamžäh* from “scrape,” FAN *tgänäh* from “barrack,” FAN *tkäsräh* from “go down,” FAN *t(ä)nändräh* from “burn,” FAN *tnändrah* from “steal,” FAN *tärḏäh* from “bathe,” FAN *ətžäwgäh* from “exile,” FAN *täžyäh* from “be quick,” and FAN *tuʔvräh* from “fill.”

- (ii) Examples of FANs in *-ih* are FAN *tīḍgih* from “be wet; (get) wet,” FAN *tīḍrih* from “mention,” FAN *tīʔnwiḥ* from “tie up; repair,” FAN *tīʔršiḥ* from “slit the throat,” UFAN *tuʔšbiḥ* from “take a sip; a puff,” FANs *təDriḥ/təd(d)riḥ* from “prick,” FANs *təgriḥ/tāguʔriḥ* from “pull up the delou; draw water from a well,” FAN *taḍuʔriḥ* from “fall down; raid,” FAN *tāš(š)əddih* from “become thin,” and UFAN *tānḍiywiḥ* from “bury.”

The patterns of all these FANs without suffix *-t* and with final *-h* vary, but the absence of a vowel between R^2 and R^3 is regular (as between R^1 and R^2 as in *tārḍāḥ*, *tāẓyāḥ*, *tār wāḥ*, and *təgriḥ* when the root has only two radicals – except for *təgānāḥ*): there is rearrangement of the scheme in relation to FANs in *-t*. This could explain the doublets *tār wāḥ/tār ukt* (FANs from “put down”), *tāẓəgrāḥ/tāẓəgārt* (FANs from “go out”), *ṭḥat(t)šāḥ/ṭḥatāš-t(i)yaṭš* (FANs from “be cut off; cut”), *ṭziʔwraḥ/tāẓaʔt* (“be or become fat”), and *tīʔgrāḥ-tīʔgriḥ/tiʔgārt* (FANs from “fly”).

Among the FANs of this type (without *-t*, but with an *-āḥ/-ih*), some arise from roots with a radical final laryngeal. In the case of H or H*, one may hesitate between the absence of the radical or the absence of the feminine ending (cf. FANs *tīriḥ/tiʔriḥ* from “dictate,” FAN *taḍāḥ* from “pant,” FAN *tuyāḥ* from “search” (*āllih*), FAN *tugāḥ* from “carry; be pregnant”). But in the case of the glottal ʔ, there is clear shortening of the radical (cf. UFAN *tīkiḥ* from “give,” FAN *təẓəššiḥ* from “drink,” FAN *tīʔnwiḥ* from “kill”). This shortening is visible with annexation: compare the MANs (with ʔ) *əẓəššiʔ-n* . . . “drinking from. . .,” *iʔḍiʔ-n* . . . “striking with. . .,” and the UFANs (without ʔ) *təẓəšši-n* . . . “the drinking (once) of. . .,” *tīʔḍi-n* . . . “striking (once) with. . .”

Conclusion

ANs are probably the least regular domain of derivational morphology. In Zenaga, the situation is all the more complex as basic verbs (generally bisyllabic) often have several ANs, sometimes of the same gender – sometimes of both genders. There are cases of specialization as the form of ANs can vary depending on the meanings of the verb. Moreover, FANs can be unitary nouns, but oppositions with MANs are rare and not always well established – the situation seems to be comparable in Tuareg (see Prasse 1974: 85). In total, MANs are much more numerous, and their most frequent patterns show “non-*a*” thematic vowels: “non-*a*—non-*a*” for bisyllabic verbs and “non-*a*—non-*a*—non-*a*” for trisyllabic and derivative verbs (except for those bearing the prefix *aʔ*).

In Berber, the “non-*a*” patterns regularly correspond to the patterns of negative verbal forms (especially the NP of bisyllabic verbs). They may, thus, appear as the basic – unmarked – vowel patterns of Berber. While the prevalence of this pattern for ANs (especially MANs) is clear in Zenaga, this is not often the case elsewhere. Heath (2005: 517), however, gives AN *udəḥ* from *əddəḥ* “pound” as an example of one of the three main types for verbs with a consonant final radical. Kossmann (2013: 87–88) notes the frequency of ANs in Ghadames with an *a*- marker and

vowels in ə (e.g., AN *akənnəf* from “roast”) and stresses the proximity to the negative IA pattern (*əkənnəf*). The *a*-marker is reminiscent of the Zenaga *a*², except that the presence of this *a*- is not limited in Ghadamsi to the AN of certain trisyllabic or derivative verbs (nor among the Aït Seghrouchen, see Bentolila (1981: 398–403)).

In Zenaga, ANs corresponding to conjugated adjectives are regularly FANs in “*a*—*a*.” In Tashlhit, Galand (2002: 228) also notes the specificity of ANs from verbs expressing qualities: FANs with initial *t*- and final *-i* (e.g., AN *taluyyi* from *ilwiḡ* “become soft”). The final vowel may correspond to the *-t*-less endings observed in Zenaga FANs and reported elsewhere, notably by Heath and Bentolila.

Vowel patterns in “*a*” are found for verbs other than quality verbs, especially for bisyllabic verbs, which form their IA by means other than prefixation of *tt*- (by doubling the radical, inserting a vowel, prefixing *y* or *n*). The existence of ANs with an IA pattern (in “*a*”) alongside ANs with a PN pattern (in “non-*a*”) is reminiscent of Prasse’s (1974: 83) distinction between “perfective” and “imperfective” ANs, but the similarity of the latter, in Zenaga, with ANs expressing a quality, makes it difficult to classify them as themes with dynamic content, as opposed to the former, which are static.

For roots with a glottal final, one observes in Zenaga that the corresponding AN bears a final *-i*. This is convergent with other Berber languages, for example, in Ghadamsi the ANs of verbs with alternating vowels like *alassi* AN of “wear” (Kossmann 2013: 88). This confirms the non-negligible, but complex, role of radicals in NA formation.

Bibliography

- Attayoub AK (2001) La tatsəret des Aytawari Seslem: Identification socio-linguistique d’un parler berbère non-documenté chez les touaregs de l’Azawagh (Niger). INALCO, Paris
- Bentolila F (1981) Grammaire fonctionnelle d’un parler berbère: Aït Seghrouchen d’Oum Jeniba (Maroc). SELAF, Paris
- Cohen D, Taine-Cheikh C (2000) À propos du zénaga. Vocalisme et morphologie verbale en berbère. Bulletin de la Société Linguistique de Paris XCV((1)):269–322
- Galand L (2002) La problématique du nom verbal en berbère. In: Nait-Zarrad K (ed) Articles de linguistique berbère. Mémoires Werner Vycichl. L’Harmattan, Paris, pp 219–234
- Heath J (2005) A grammar of Tamashek (Tuareg of Mali). Mouton, Berlin
- Kossmann MG (1999) Essai sur la phonologie du proto-berbère. Rüdiger Köppe, Köln
- Kossmann MG (2013) A grammatical sketch of Ghadames Berber (Libya). Rüdiger Köppe, Köln
- Mourigh K (2016) A grammar of Ghomara Berber (North-West Morocco). Rüdiger Köppe, Köln
- Nait-Zerrad K (1998–2002) Dictionnaire des racines berbères (formes attestées) (3 vols.). Peeters, Louven
- Prasse K-G (1974) Manuel de Grammaire Touaregue (tāhāggart) IV–V Nom. Editions de l’Université de Copenhague, Copenhague
- Prasse K-G, Alojaly G, Mohamed G (2003) Dictionnaire Touareg-Français (Niger). Museum Tusculanum Press–Université de Copenhague, Copenhague
- Taifi M (2006) Syntaxe et sémantique du nom verbal en tamazight: cas du complément interne. In: Vossen R, Ibrizimow D, Stroemer H (eds) Études berbères III. Le nom, le pronom et autres articles. Rüdiger Köppe, Köln, pp 239–251

- Taine-Cheikh C (2002) Morphologie et morphogenèse du diminutif en zénaga (berbère de Mauritanie). In: Naït-Zerrad K (ed) *Articles de linguistique berbère. Mémorial Werner Vycichl*. L'Harmattan, Paris, pp 427–454
- Taine-Cheikh C (2003) La corrélation de gémation consonantique en zénaga (berbère de Mauritanie). *Comptes rendus du GLECS* 34(1998–2002):5–66
- Taine-Cheikh C (2009) À propos de l'expression de l'état en zénaga. Apophonie et sous-catégorisation verbale en berbère et en arabe. *Studi Mayribini* (N. S.) 7:95–109
- Taine-Cheikh C (2012) L'expression du haut degré en zénaga. Étude sémantique et morphosyntaxique. In: Ibrizimow D, Vossen R, Stroomer H (eds) *Études berbères VI. Essais sur la syntaxe et autres articles*. Rüdiger Köppe, Köln, pp 269–289
- Taine-Cheikh C (2014) Qualification and comparison in Berber. The verb-noun distinction and its fluctuations. *STUF* 67(1):63–79
- Taine-Cheikh C (2015) La ronde des saisons. . . et des labiales! Des noms des saisons en berbère. *Studi Africanisti. Quaderni di Studi Berberi e Libico-Berberi* 4:297–319
- Taine-Cheikh C (2019) Les préformantes *m* et *n* en zénaga. Dérivations et significations. In: Vossen R, Winkelmann K, Ibrizimow D, Stroomer H (eds) *Études berbères VII. Essais sur la sémantique en berbère et autres articles*. Rüdiger Köppe, Cologne, pp 189–218

Section V

Lexicology and Onomastics

A Preliminary Account of Hypocoristics in Moroccan Amazigh

31

Abdelaziz Boudlal and Mohamed Yeou

Contents

Introduction	662
Review of the Literature	664
Hypocoristics in Moroccan Amazigh	667
Deriving Hypocoristics in MAM: A Preliminary Account	671
Pattern I	673
Pattern II	674
Root-Based Hypocoristic Patterns	674
Variation as Reflective of a Socio-pragmatic Function	677
Conclusion	678
Appendix	679
Hypocoristics in MAM	679
References	681

Abstract

This paper purports to undertake a taxonomic survey of hypocoristics in one of the Moroccan Amazigh varieties, namely, the variety spoken in Figuig, and described in Yeou (Dictionnaire Amazighe-Français. Parler de Figuig. Publications de

This chapter is a modified version of an earlier conference paper by the same authors, which appeared as Boudlal and Yeou (2019). The authors would like to thank the editors for having accepted to publish this chapter, which has tremendously benefited from comments and suggestions made by two anonymous reviewers, leading to its improvements. Any remaining errors or inconsistencies are our entire responsibility. Finally, the authors' names appear in alphabetic order.

A. Boudlal

Faculty of Letters and Humanities Ben M'sik, Hassan II University of Casablanca, Casablanca, Morocco

e-mail: abdelaziz.boudlal@univh2c.ma

M. Yeou (✉)

Faculty of Letters and Humanities, Chouaib Doukkali University, El Jadida, Morocco

e-mail: yeou.m@ucd.ac.ma

l'IRCAM, Rabat, 2022). The paper is also a contribution to the ongoing debate on root-based – as opposed to word-based morphology. In particular, it considers whether or not hypocoristics take the root, or the word, as the base of derivation. The paper tries to show that the derivation of hypocoristics in Moroccan Amazigh is better analyzed as a case taking both the root and the word as a base. Facts from the language lend support to this assumption: hypocoristic nouns such as *kari*, *ttu*, and *jja* (from *ʕabdəlkarim*, *kəttu*, and *məkkijja*, respectively) should, in principle, take the word as a base, while nouns such as *mənna*, *ħəffu*, and *basa* (from *jamina*, *məħfud*¹, and *bənnas*¹*ər*¹, respectively) take the root, or part of it, as a base to derive disyllabic hypocoristics. Both root- and word-based hypocoristics are subject to a truncatory process which yields an output that corresponds to a minimal word. The paper also provides a preliminary account of a number of hypocoristics within the constraint-based framework of optimality theory.

Keywords

Word formation · Root · Word · Hypocoristic · Truncation · Moroccan Amazigh · Figuig

Introduction

Hypocoristic forms of personal names are used in a language to signal a type of relationship between the speaker and the addressee when they engage in a conversation. For Newman and Ahmad (1992: 159), hypocoristic forms in some languages such as Hausa and English may both portray both the affection of the speaker and the diminutive nature of the referent. For Schneider (2003: 145), they “signal social nearness in an informal communication situation. They are used between equals, but not for addressing children.”

The present study describes hypocoristics in the variety of Moroccan Amazigh¹ (MAM) described by the second author in a dictionary (2022), titled *Dictionnaire Amazigh-Français. Parler de Figuig*, in which more than 100 hypocoristics, were compiled. In this variety, hypocoristics constitute the standard way of addressing adults. They relate to a positive social aspect, which is to show respect to the referent. In the variety described, hypocoristics are considered to be status markers. A younger speaker should never address an older person by his full name; instead, he should use an appropriate hypocoristic form used for adults. The older person usually addresses a younger person by his full name or sometimes by a hypocoristic form for young people, when available. The latter hypocoristic form is also used by young people of the same age or the same generation.

In a study on the function of hypocoristics as used in Figuig, the anthropologist Souad Azizi considers them as a mechanism allowing for the amazighization of Arab-Muslim full names (Azizi 2013). Such full names are encoded in writing for the purpose of

¹The variety of Moroccan Amazigh spoken in Figuig has been subject to a number of studies (e.g., Ben-Abbas 2003; Benamara 2011, 2013; Kossmann 1997; Saa 2010).

registration in the State's civil system or for documenting legal and financial acts. However, in communicative contexts, hypocoristic forms are the ones mostly used by people to identify themselves. In this system of identification, the standard practice is to name someone by his first name followed by his/her family name, which is usually preceded by a particle denoting his/her patronym or father's lineage. This particle has two forms: /u/ is used for males and /ut/ for females. Illustrative examples are given in (1):

- (1) *mmu u ħəmmuʃ* Mou (Mohamed) Hammouche
tatta ut bəlʕid Tatta (Fatima) Belaïde

When people have the same first name as well as the same family name, it is common to identify them by referring to their father's or their mother's first name; males by referring to their father's first name, and females by their mother's name. In this case, the genitive particle /n/ precedes the father's and the mother's first names, as illustrated in (2):

- (2) *mmu n dudḡu u* Mou (Mohamed) n (son of) dudḡu (Ahmed)
ħəmmuʃ Hammouche
tatta n ʃʃa ut bəhi Tatta (Fatima) n (daughter of) ʃʃa (Aïcha) Bahi
mmu n daḡḡa Mou (Mohamed) n (son of) Hadda

It should be noted that in fast speech, the particles /u/, /ut/, and /n/ pronounced in careful speech may be deleted, resulting in a sequence of two or three hypocoristic forms, as in *mmu dudḡu ħəmmuʃ*.

As for a married woman, she may be identified by using the genitive particle /n/ followed either by the name of her father or her husband or by the particle /at/ followed by the name of her husband's patronym:

- (3) *ʕʕa n buʕʕa* Âa (Jomâa) n (wife of) Bouâa (Boujemâa)
nna n mmu Nna (Amina) n (daughter of) Mou (Mohamed) ʕarabi
ʕarabi (Arabi)
tatta n-at jəʕʕu Tatta (Fatima) n-at (of the family) Yaôu

This paper surveys the different patterns of hypocoristics in MAm (also referred to as shortened personal names) and addresses two main questions. The first question relates to the nature of truncation hypocoristics undergo in terms of word structure; that is the constraints regulating this truncation and whether or not it is governed by any minimality requirement on their output. The second question concerns the nature of the input that serves as a base for the derivation of hypocoristics: is it the root, or is it the word? It should be made clear here that an exhaustive analysis of hypocoristic derivation is beyond the scope of this paper. Nevertheless, we will point out to directions on how they can be dealt with, suggesting the optimality-theoretic (OT) model proposed in Prince and Smolensky (1993/2004) and McCarthy and Prince (1993).

The rest of the paper is articulated as follows. Section “[Review of the Literature](#)” will review a select body of literature about hypocoristics in other languages. Special attention will be given to languages such as Italian, Spanish, Arabic dialects, and English.

Section “[Hypocoristics in Moroccan Amazigh](#)” will present the data and provide a description and patterning of hypocoristics in MAm. Section “[Deriving Hypocoristics in MAm: A Preliminary Account](#)” tries to capture generalizations about this category of nouns by recourse to the nature of the base serving as input. In particular, we provide a constraint-based account, couched within the OT framework, pointing out to directions of what a preliminary account of hypocoristics looks like. Section “[Variation as Reflective of a Socio-pragmatic Function](#)” will consider the socio-pragmatic status of hypocoristics; and finally, Section “[Conclusion](#)” culminates our findings.

Review of the Literature

This section is meant to review the literature on some of the cases dealing with hypocoristics across select languages. Hypocoristics often undergo a process whereby part of the name is truncated. By truncation we mean “a process where a morphological category is typically marked through lack of material in the derived form as compared to the base form” (Alber and Arndt-Lappe 2012: 310). Most often, this truncatory process is regulated by the prosody, and the output is subject to a constraint on the structure of the output. Depending on the language under study, the output often corresponds to the minimal word of the language concerned, which is bimoraic or disyllabic, depending on whether the language in question is quantity-sensitive or insensitive. In this respect, Alber and Arndt-Lappe (2012: 290) distinguish two types of truncation: templatic and subtractive. In templatic truncation, the size of the derived form is predictable, whereas in subtractive truncation, it is the size of the truncated material which is predictable. The MAm truncation cases considered in this paper relate to templatic truncation.

The hypocoristic examples we will consider in this section relate to Italian (Alber 2010), Spanish (Álvarez 2015), Colloquial Arabic (Abu-Mansur 2019), and English (Benua 1995). The objective is to see how the derivation of hypocoristics in these languages can relate to the cases of MAm we will consider.

Alber (2010) studies truncation in a productive pattern of hypocoristics in Italian. She distinguishes two disyllabic patterns which preserve material at the left edge of the name base as the examples in (4) below show:

- (4) a. Anchoring to the left edge of the name base (Alber 2010: 2):
- | | |
|--------|-------------------------|
| Fránce | Francésca |
| Vále | Valentína / Valentíno |
| Ále | Alessándra / Alessándro |
- b. Anchoring to the left edge with final [i]:
- | | |
|--------|-----------|
| Francy | Francesca |
| Andri | Andrea |
| Steffi | Stefania |

In (4a), truncation yields a disyllabic output which consists of material from the first segment to the second vowel of the input. Pattern II (i.e., 4b) replaces the second vowel by the default vowel *i* in Italian.

Other patterns in Italian seem to abide by stress factors and, therefore, favor an output form that preserves the stressed vowel as shown by the following examples from Alber (2010: 3):

(5) c. Anchoring to the stressed syllable of the name base:

Césca	Francésca
Bérto	Robérto
Nóra	Eleonóra

A final pattern of hypocoristics in Italian preserves the stressed syllable of the full name and reduplicates the onset following the base stressed syllable, as shown in (6) below:

(6) Anchoring to the stressed syllable and reduplication:

Pípipo	Filípipo
Péppe	Giuséppe
Gígi	Luígi

Another language that makes recourse to a similar type of truncation is Spanish, as described in Álvarez (2015). This language exhibits three types of hypocoristics: left-anchored, stress-anchored, and reduplicative. Let us consider the first type of truncation in (7) below:

(7) Left-anchored truncation (Álvarez 2015: 12):

<i>Base form</i>	<i>Nickname</i>
Manu[é]l	Manu
Rafa[é]l	Rafa
Ter[é]sa	Tere
Ver[ó]nica	Vero

In these examples, the retained syllables are the stressed ones and the syllable immediately preceding it.

The second type of truncation hypocoristics undergo is also stress-anchored; here these hypocoristics truncate all but the stressed syllable and the syllable that follow, as shown in (8) below:

(8) Instances of stress-anchored truncation (Álvarez 2015: 15):²

<i>Base form</i>	<i>Nickname</i>
Alb[é]rto	Be(r)to
Alf[ó]nso	Poncho
Asun[θ]ión	[tʃ]on-a
Aur[ó]ra	Lola

²Note here that the words presented in (8) exhibit alternations that are not directly relevant to the main point made. These concern the alternations *f/p*, *θ/tʃ*, and *r/l* affecting the last three words of the data set.

The third type is reduplicative and seems to duplicate the consonant serving as the onset of the hypocoristic word as the examples in (9), below, show:

(9) Some instances of reduplicative truncation (Álvarez 2015: 21):

<i>Base form</i>	<i>Nickname</i>
Ana	N-ana
Bonifacio	[tʃ]-a[tʃ]o
Carlos	Ca-c-o
Carlota	T-ota

In dialectal Arabic hypocoristics have received little attention, albeit there are two published works. The first, by Davis and Zawaydeh (1999), is a descriptive analysis of hypocoristics in Ammani Jordanian Arabic. The second reference, elaborated within a theoretical framework, is that of Abu-Mansur (2019). The author assumes that there are two main patterns of hypocoristics in Makkan Arabic. Representative examples are given in (10) below:

(10) a. Pattern I hypocoristics (Abu Mansur 2019: 32):

hasan	hassuun
muhəmməd	hammuuda
ʔiʕtidaal	ʕadduul(a)

b. Pattern II (Abu Mansur 2019: 33):

faza	fazzu
zakiyyah	zakku
nada	naddu

Within the constraint-based OT framework, Abu Mansur (2019) provides an analysis which shows that syllable structure constraints and the OCP account for the apparent differences between these two hypocoristic forms. The author holds that in order to understand the formation of hypocoristics in this Arabic dialect, reference needs to be made to, both the lexical and output root – a fact which will be shown to be also operative in MAm.

Another Arabic dialect that shows commonalities with Makkan Arabic in the formation of hypocoristics is Moroccan Arabic. In this variety, hypocoristics are governed by a prosodic template requiring that their output conforms to a disyllabic word on the pattern CəC_iC_iuC, as shown in the examples we provide below:

(11) Hypocoristics in Moroccan Arabic:

	Full name	Hypocoristic
a.	məsʕʔafa	sʕʔʕʔuf
	ʕaziz	ʕəzzuz
	xalid	xəllud
b.	xadiza	xədduʒ
	fatʕima	fəʕʕʕum(a)
	naʕima	nəʕʕʕum(a)

The items in (11a) are male names, whereas those in (11b) are female names, a fact which explains why the second set sometimes adds the feminine suffix *-a*. These hypocoristics are derived by taking the consonantal root and mapping it onto the template CəC_iC_iuC. Note here that it is the second consonant of the triconsonantal root that consistently reduplicates.

The final case of hypocoristics we will consider in this section comes from English, and more particularly certain dialects of New York and Philadelphia, studied in Benua (1995). In these examples, the truncated names in both (12a) and (12b) preserve the vowel quality of their corresponding full form. Normally, the English low front vowel [æ] is tensed to [a] in syllables closed by an [r] which precedes another consonant or a pause. This æ-tensing never takes place in hypocoristics as the examples below show:

- (12) a. Benua (1995: 2)
- | | | | |
|-------|-------|-----|-----|
| hæ.ri | Harry | hær | Har |
| læ.ri | Larry | lær | Lar |
| sæ.rə | Sarah | sær | Sar |
- b. Benua (1995: 9)
- | | | | |
|----------|--------|-----|-----|
| jæ.nIs | Janice | jæn | Jan |
| pæ.mə.lə | Pamela | pæm | Pam |

Forms showing tense vowels (i.e., **har*, **lar*, **sar*, **jEn*, **pEm*) never surface because of the nature of the relationship regulating the full and the truncated names. The non-tensing of the vowel [æ], forced by the constraint **[ær]*, has driven Benua (1995) to conclude that truncated words are related to the surface form of their sources since the truncated words faithfully reproduce derived surface properties of the source words.

In sum, the hypocoristics considered in this section have a common denominator. They are all regulated by a template imposing a certain limit on the output form: most often disyllabic or bimoraic (as is the case with English). Except for the Arabic data in (10) and (11), the other hypocoristics resort to truncation in order to achieve this minimality requirement. The material truncated is regulated by constraints which specify which part of the name is truncated and which is retained. The output can thus be left-anchored, stress-anchored, or both.

In what follows, we will present the data exemplifying cases of hypocoristics in MAM. Whenever we deem it relevant, we will try to draw a comparison between MAM and the languages exemplified above, both in terms of word structure output and also the truncatory phenomenon exhibited.

Hypocoristics in Moroccan Amazigh

As stated earlier, the data in this study are based on Yeou (2022). We have been able to group the hypocoristics in terms of five categories we will be referring to as patterns. Our classification is based on their behavior with respect to two main criteria: first, the part of the name that is truncated, and, second, the morphological

form that serves as the base of derivation for the hypocoristic. One generalization to make about hypocoristics in MAm is that they all involve some kind of truncation.

The first type we will consider is one where the right-hand part of the base noun is preserved and the left-hand side truncated, resulting in a disyllabic output, which we refer to as right-anchored Pattern I hypocoristics:

(13) Pattern I: Right-anchored *CVCV* hypocoristics³

Source name	Hypocoristic	
r ^ʕ abiʕa	biʕa	Rabiâa
lmadani	dani	El Madani
məʕzuza	zuza	Maâzouza
kəzzaza	zaza	Kezzaza
məns ^ʕ ur ^ʕ a	sula	Mansoura
ʕabdəlkərim	kari	Abdelkarim

These shortened names tend to target the last syllable of the source word. In the case of a compound word such as ʕabdəlkərim, the first element (i.e., the word ʕabd) is wholly truncated. Also note the deletion of the coda in the hypocoristic (i.e., the consonant *m*), a fact which aligns this name with the other hypocoristics which generally tend to end up with an open syllable.

Pattern II of hypocoristics we will consider are items that are also right-anchored but whose output form has resorted to the gemination of the final base consonant, giving rise to Pattern II hypocoristics with an initial geminate, followed by a vowel, as shown by the items below:

(14) Pattern II: Right-anchored *C_iC_iV* hypocoristics with initial gemination:

Source name	Hypocoristic	
bufəlʒa	ʒʒa	Boufelja
ʒəmʕa	ʕʕa	Jemâa
ʃixa	xxa	Chikha
ʕi(j)ʕa	ʃʃa	Aïcha
məs ^ʕ tʕafa	fʔa	Mustafa
xadiʒa	ʒʒa	Khadija

One legitimate question to ask here about these items is why they do not form hypocoristics in the same way the items in (13) above do; in other words, why do items such as *bufəlʒa* and *xadiʒa* not truncate to surface as *ʔəlʒa and *diʒa instead of the expected form ʒʒa? We do not have any definite answer to this question right now but suffice it here to mention that only the final syllable of the input noun is retained, and because this syllable does not meet the minimality requirement on the

³Note that the sound *r* changes to *l* in the truncated form, something that we are unable to explain as there are similar items such as ʕabdəlkərim with the same sound that does not change when truncated (*kari*).

output form (disyllabicity), recourse is made to the gemination of the onset of the last syllable.⁴ A tentative explanation would be to assume that truncation in (14) operates on unsyllabified roots and that only the last syllable is maintained. In this case, gemination will have to apply in order for the output to conform to word minimality requirement. However, this solution would produce unattested forms such as **fafa*, **diza*, and **lla* instead of the correct hypocoristics *ffa*, *zza*, and *lla* from the corresponding names *mās^hafa*, *xadiza*, and *buɣlam*.

Pattern III of hypocoristics we will consider in this paper will be referred to as reduplicative Pattern III forms, since these names geminate one of the base segments. The base here happens to be one of the root segments. These hypocoristics abide by a predefined pattern CVC_iC_iV and always present a geminated consonant in the middle. Which root consonant is geminated is unpredictable. It can be C1 (e.g., *zazza* in (15a)), C2 (e.g., *tatta* in (15a)), or C3 (e.g., *sidḡa* in (15b)):

(15) Pattern III: CVC_iC_iV hypocoristics with medial gemination:^{5,6,7}

	Source name	Hypocoristic	
a.	zzəhr ^h a	zazza	Zahra
	fat ^h ima	tatta	Fatima
	mbar ^h ka	bakka	M'barka
	ʒəmɣa	ʒaɣɣa	Jemâa
	buzijjan	bazza	Bouziane
b.	saɣid	sidḡa	Saïd
	xadiza	xizza	Khadija
	bəlqasəm	baqqu	Belqassem
	ħməd	dudḡu	Ahmed

⁴It seems that Pattern II C_iC_iV hypocoristics correspond to the minimal word only if we consider the initial consonant to form a minor syllable much like the case of Moroccan Arabic words such as *gra* “study,” and *fra* “to buy,” where the first consonant forms a minor syllable, and the second consonant forms a major syllable along with the vowel *a* (see Boudlal 2001, for more details on the difference between major and minor syllables).

⁵At a first glance, it seems that the deparyngealization of emphatic sounds applies to hypocoristics to avoid marked structure, something that is further corroborated by examples such as *bənnas^hər* “Bennacer”/basa, where *s^h* deparyngealizes. However, upon considering other hypocoristics, we realize that when the emphatic sound is part of the truncated portion of the base, the feature pharyngeal is docked onto the neighboring sound as is the case with names such as *br^hahim* “Brahim” and *ɣumar^h* “Omar” truncated to *b^hab^ha* and *ɣəm^hm^ha*, respectively. In sum, deparyngealization in the variety of MAm considered here is unpredictable and as such very hard to capture in terms of a context-specific rule.

⁶Notice that the name “Khadija” has two hypocoristics *xizza* and *zza*. The choice of one of the two forms is decided on the basis of socio-pragmatic functions. See section “[Variation as Reflective of a Socio-pragmatic Function](#)” below for more details.

⁷Notice here the devoicing of the final consonant of the base *d*, a characteristic phenomenon of the oral stops /b, d, g/, when they geminate. The voiced geminates /bb, dd, dd^h, gg, gg^w/, which Figuig shares with other Amazigh varieties, are realized as voiceless both at the lexical and morphophonemic levels. The devoiced coronal geminate, which is produced without aspiration is transcribed as [d̥d̥] to distinguish it from the voiceless aspirated geminate [tʰtʰ] (Yeou et al. 2011).

Note here that the organization of the data sets in (15) is based on the vocalic pattern of the hypocoristics. In (15a), the pattern is *a_a*, which is the more productive and recurrent pattern – as opposed to those in (15b) which show variation between *i_a*, *a_u* and else *u_u*.

A fourth *CVC_iC_iV* pattern of hypocoristics, exhibiting medial gemination of a root segment, is the one with schwa and a full vowel, where the vowel could be *i*, *a*, or *u*. For illustration, consider examples of what we label Pattern IV hypocoristics in (16) below:

(16) Pattern IV: *CəC_iC_iV* hypocoristics with medial gemination:

Source name	Hypocoristic	
bəxta	xətta	Bakhta
bəlxir	bəxxi	Belkhir
məħfud ^ə	ħəffu	Mahfoud
buʕlam	ʕənnna	Bouâalam

It seems that the derived hypocoristic in (16) keeps the second vowel of the input and then makes recourse to schwa epenthesis to satisfy the disyllabic requirement on the output. The result is a word on the pattern *CəC_iC_iV*.

The final type of hypocoristics found in MAM, which we will be referring to as Pattern V, is of a mixed type which may or may not show gemination but abides by a template consisting of two light syllables whose vocalic pattern shows variation as the examples in (17) show:

(17) Pattern V: Mixed *CVCV* hypocoristics:

Source name	Hypocoristic	
mulud	lulu	Mouloud
sliman	sima	Slimane
ʕabdləmʒid	ʒida	Abdelmajid
muhəmməd	muda / madi	Mohamed
br ^ə ahim	bahu	Brahim
bəlqasəm	qasu	Belkacem
ʕabdəlwfafi	fafi	Abdelwafi

The mixed nature of these hypocoristics stems from the fact that it is impossible to characterize them in terms of their consonantal or vocalic constituency. They could show right alignment of the base form (*bəlqasəm/qasu*), right alignment of the base form followed by final consonant reduplication (*ʕabdəlwfafi/fafi*), left alignment of the base (*sliman/sima*), initial reduplication of a base consonant (*mulud/lulu*), or any such mixed type (*muhəmməd/muda*). The vocalic patterns are also of a mixed nature: they could be *u_u*, *i_a*, *u_a*, *a_i*, or *a_u*.

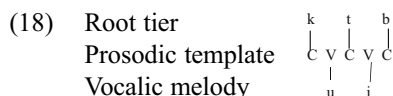
Looking back at the data presented in this section, a number of generalizations could be made in relation to our MAM hypocoristics. First, they are all subject to a compulsory process that truncates part of the base. Second, the output resulting from this truncation retains the right-hand portion of the base in some of the cases. Third,

MAM, and the rest of the languages reviewed, seems to abide by a prosodic constraint limiting the size of the hypocoristic. Most of the examples considered are either disyllabic or bimoraic. Finally, MAM seems to differ from the other languages considered in one derivational aspect. While the derivation of hypocoristics seems to be predictable in all the languages considered, we have noticed that a substantial body of data from MAM seems to pose a real derivational problem, especially relating to the root consonants involved. This makes it harder to come up with an analysis that could be generalizable to all the hypocoristics in the language, representative examples of which are given in the appendix.

Deriving Hypocoristics in MAM: A Preliminary Account

This section provides a preliminary account of hypocoristic formation in MAM. Our classification of the data in the previous section has been established on the basis of the nature of the input to hypocoristic derivation. The main assumption underlying this work is that hypocoristics may take as input either the root or the word. In order to better understand the ongoing debate on root-based versus word-based morphology, a short review of the concept of the root is in order.

Within the prosodic theory of non-concatenative morphology – a framework proposed in McCarthy (1981) to deal with word formation processes in Semitic languages – a word is seen to consist of three different morphemes, represented on different levels, referred to as tiers: the consonantal root tier, which carries the basic meaning, the vocalic melody tier, and the prosodic template tier. A word such as *kutib* “write (perfective, passive)” will be represented as in (18) below:



This root-based approach has also been applied to the morphology of MAM, particularly in the compilation of dictionaries such as Oussikoum (2013), Serhoual (2002), Taifi (1992, 2016), and Yeou (2022). However, McCarthy’s (1981) traditional concept of the root has been challenged in a number of works such as Bat-El (1994), Ussishkin (1999), Benmamoun (1999), Schluter (2013), Bensoukas (2018), and Boudlal (2018), among others, all of which argue that the stem – or the word can also serve as a base for word formation. For example, Bensoukas (2018) argues against the consonantal root in Tashlhit (a variety of Moroccan Amazigh). In Boudlal (2018), the view that the root in MA consists, mainly, of abstract consonants is discarded, and a new definition of the “root,” based on previous works by Heath (1987) and Al Ghadi (1990/2014), is provided to the effect that the “root,” which will be referred to here as “base,” is seen to consist of both consonants and vowels.

With this background in mind, we turn to hypocoristics in MAM to see the nature of the base that serves as an input to their derivation. To do so, let us recapitulate some of the generalizations that we have made above. First, types I and II

hypocoristics are all anchored to the right edge of the prosodic word. Second, they are all constrained by a prosodic template requiring that they be exactly bimoraic. Except for type II, where the hypocoristic word consists of a minor syllable, followed by a major one (Boudlal 2001), all other types consist of two major syllables.⁸ Patterns III and IV show reduplication of one of the consonants of the base, and Pattern V is a mixed type, consisting of a sequence of light syllables with or without root consonant reduplication. Finally, all hypocoristics end up in an open syllable, a fact which shows that syllable codas cannot be word final. One could possibly argue that codas are constrained in words such as *zazza*, *xətta*, and *ʕəzzu* where the final consonant is geminated since they have the same feature as the onset of the final syllable. However, the existence of hypocoristics such as *baħu*, *muda*, and *zida* (from *ʕabdəlmazid*, *muħəmmad*, and *br^ʕahim*) is a striking evidence that codas are not allowed word finally.

To try to account for hypocoristics in MAm, let us refer to Standard OT (Prince and Smolensky 1993/2004; McCarthy and Prince 1993, and related works), one of the frameworks that has had a lot of success in recent years. OT is a theory of constraints and their interaction. Its apparatus consists of three elements: a generator (GEN), an evaluator (EVAL), and a set of universal constraints (CON). GEN generates output candidates which are evaluated by EVAL based on how well they satisfy CON. For illustration, consider the constraint tableau below, where CON consists of Constraints A and B, in this respective order:

(19)

Candidates	Constraint A	Constraint B
☞ a. Candidate 1		*
b. Candidate 2	*!	

The optimal candidate is marked by ☞. Violations are marked by an asterisk [*], and a fatal violation is marked by [*!]. Constraint A dominates Constraint B (written as Con A >> Con B), represented in the tableau by a solid line separating the two constraints. When no domination holds between the two constraints, Con A and Con B are written as Con A and Con B and are separated by a dotted line in the constraint tableau. The optimal candidate in this tableau is candidate A which violates a low ranked constraint, namely, Con B.

In what follows, we provide a preliminary account of hypocoristics in MAm couched within the OT framework. The data in (13–17) above will be divided into two major types: hypocoristics that take the word as a base (Patterns I and II) and those that take the root (Patterns III, IV, and V).

⁸Boudlal (2001) makes a distinction between a major and a minor syllable. A major syllable is a syllable headed by one of the full vowels /i, u, a/, or schwa, and a following consonant. A minor syllable is one consisting of a single consonant, dominated by a mora.

Pattern I

Pattern I hypocoristics are characterized by a truncation of part of the base, which we assume to be the word and not the root. Words such as *biša* and *dani* (from *rʔabiša* and *lmadani*, respectively) are the result of a truncatory process whereby the left-hand side of the word is truncated to yield an output consisting of a disyllabic word. The main constraints at play, here, are given below:

(20)

ONSET (ONS)

Onsetless syllables are banned

FOOT BINARITY (FT-BIN)

Feet are binary under syllabic or moraic analysis

ANCHOR- RIGHT (base/truncated) (ANCHOR-R)

The right edge of the truncated form must be aligned with the right edge of the base

MAXIMALITY (base/truncated) (MAX-BT)

All the segments of the base correspond to the segments of the truncated form

The constraint ONS requires that syllables in hypocoristics begin with a consonant. FT-BIN ensures that the output form corresponds to a prosodic word. It acts as a prosodic constraint requiring that the hypocoristic form should consist of at least two moras or syllables. The second constraint, that is, ANCHOR-R, specifies that part of the base which is retained. It corresponds to the right edge of the prosodic word. The third constraint (MAX-BT) bans any deletion of base material. Given that hypocoristics are the result of a truncatory phenomenon, it follows that the constraint MAX-BT must be low-ranked. ONS is undominated and as such will not figure in the tableaux below. The constraints FT-BIN and ANCHOR-R, which should not be ranked with respect to each other, are observed in hypocoristics and are, therefore, high-ranked as the tableau below for the base *rʔabiša* shows:

(21)

Base: rʔabiša	FT-BIN	ANCHOR-R	MAX-BT
ʔa- biša			**
b- rʔabi		*!	**
c- ša	*!		****

Candidate (21b), though it observes FT-BIN, is illicit because it fails to right-anchor the base and the output. The optimal candidate spares right-anchoring at the expense of low-ranked MAX-BT.

Cases from Pattern I such as *məʃzuza* and *kəzzaza* form their hypocoristics in the same way and are governed by the constraints in (20) above. They truncate the first schwa syllable to yield a disyllabic word that is right-anchored to the prosodic work.

Pattern II

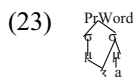
Pattern II hypocoristics are characterized by the truncation of the left-hand part of the base noun and the gemination of the final consonant. Words such as *ʒʒa* and *ʒʒa* (from *xadiʒa* and *ʒamʒa*, respectively) retain only the final syllable of the base noun as do the candidates in tableau (22) below:

(22)

Base: xadiʒa	FT-BIN	ANCHOR-R	MAX-BT
⊖a- diʒa			**
b- xadi		*!	**
c- ʒa	*!		****

The constraints we have at our disposal up to now seem to make the wrong prediction as to the optimal candidate. They wrongly predict that the optimal form is (22a), a form that satisfies all but the low-ranked constraint MAX-BT. However, as the data presented in (14) above show the optimal candidate seems to be the one where the final consonant of the noun base is geminated.

It seems that the gemination of the final consonant is dictated by the requirement that words need to be bimoraic. As we have pointed out before, the minimal word in MAM should also correspond to two moras in the same way witnessed for MA (Al Ghadi 1990/2014; Boudlal 2001). We think that the output form *ʒa* should be ruled out because it cannot sustain the required moraic structure, and as such recourse is made to the gemination of the final consonant of the base. The output obtained has a moraic structure corresponding to the form in (23):



This form satisfies word minimality by granting a moraic status to the first member of the geminate consonant, much like other varieties of Amazigh (Bensoukas 1994). Note here that word minimality can also be achieved by initial epenthesis of a vowel, thus producing hypocoristics such as **iʒa* or **aʒa*, forms that are banned by high-ranked ONS constraint.

Root-Based Hypocoristic Patterns

Hypocoristics which we assume to be based on the root are those referred to as Patterns III–V. These are characterized by a high degree of unpredictability, which makes a formal account within the OT or any other familiar framework hard to achieve. Therefore, this section will try to point to possible paths to follow if an adequate account is ever to be achieved.

The major generalizations to make about root-based hypocoristics (i.e., Patterns III–V) is that, first, they all conform to a disyllabic template, a fact we have shown they share with the word-based hypocoristics. Their output always conforms to two

major syllables, as opposed to Pattern II which consists of a minor syllable followed by a major one, as shown in (23). Second, all the patterns end up in an open syllable and confirm the hypothesis that codas are not allowed at the end of hypocoristics (see, for example, hypocoristics such as *sidḍa* and *dudḍu* from *saʕid* and *ḥməd*, respectively). Third, root-based hypocoristic patterns (except Pattern V) involve some kind of reduplication of one of the base segments. Which segment should reduplicate is the hardest to determine.

In what follows, we will show how these items could possibly be accounted for. To this end, we have tried to regroup these root-based hypocoristics into four different sets based on which of the three root consonants is geminated.

For illustration, let us consider the first set of items given in (24) below:

(24) Hypocoristics with the gemination of an unpredictable root consonant:⁹

Name	Hypocorsite	Root
zzəhr ^ʕ a	zazza	zhr ^ʕ
ʕabdəsslam	sassa	ʕbd-slm
ḥməd	dudḍu	ḥmd
rqiija	jajja	rqi
fat ^ʕ ima	tatta	ft ^ʕ m

A first look at the data gives the impression that hypocoristics favor less marked structures which consist of coronal consonants. However, that is not always the case since there are forms such as *xxa*, *ʕʕa*, and *qqa* (from *xira*, *ʕəmʕa*, and *ʕabdəlqadər*) which attest against that. What these items have in common is that they all surface with only one root consonant and that is the consonant that the hypocoristic consists of. In the first two items, the first consonant of the root reduplicates, leaving astray the other two consonants. In the case of the remaining items, it is either the final root consonant (*ḥmd* and *rqi*) or the second root consonant (*ʕʕm*). Notice further that root pharyngealization is not retained in the first and last hypocoristics, something we have shown to be unpredictable (see footnote 5 above).

The second set of root-based hypocoristics we will consider are those names that surface with two consonants, namely, the first and the second root consonants, as the examples below show:

(25) Hypocoristics with the gemination of C2:

Name	Hypocoristic	Root
məḥfud ^ʕ	ḥəffu	ḥfd ^ʕ
ḥasan/ḥusin	ḥəssi	ḥsn
ʕumar ^ʕ	ʕəm ^ʕ m ^ʕ a	ʕmr ^ʕ

⁹Note here that the root *ʕbd* is always the first member of a compound used to combine with other nouns to form personal names (e.g., *ʕabdləʕʕiz*, *ʕabdər^ʕr^ʕahim*, *ʕabdəʕʕa(h)*). In the case of hypocoristics, this first member always truncates. For more details on this truncatory process, the reader may refer to Boudlal (2001).

Notice that the first consonant in *māḥfid*⁶ is not part of the root; it stands for the Arabic passive participle marker and does not contribute in any way to the formation of the corresponding hypocoristic form. The geminated consonant in the three items is always the second consonant of the root. The third consonant never surfaces, and this might be due to the fact that hypocoristics are always vowel final; in other words, they do not allow codas at the end of a prosodic word.

The third set of hypocoristics to be considered here consists of items that also surface with only two of the three-root consonants, as shown in (26) below:

(26) Hypocoristics with the gemination of C3:

Name	Hypocoristic	Root
xadiʒa	xizʒa	xdʒ
saʕid	sidda	sʕd
xatima	xamma	xtn
jamina	mænna	jmn

In these forms, it is the third consonant of the root that is geminated, and it is this consonant that surfaces in the hypocoristic along with the first root consonant. The final word in the set is an exception since both the second and the third consonants surface, leaving the first astray.

There is still one more set of root-based hypocoristics to consider, and that is the mixed pattern in (17) above. Illustrative examples are reproduced in (27):

(27) Hypocoristics with no gemination:

Name	Hypocoristic	Root
mulud	lulu	m-wld
ʕabdləmʒid	ʒida	ʕbd-mʒd
muḥammad	muda / madi	m-ḥmd
brʕahim	bahu	brʕhm

These hypocoristics do not show any gemination and are unpredictable as to which of the root consonants should be retained and which should be truncated. For these items, we have pointed out that they are listed in the lexicon as they are, and as such the native speaker internalizes the lexical representation of both the name and its corresponding hypocoristic.

To recapitulate, this section has considered root-based hypocoristics, initially dispersed over Patterns III–V. Given their intricate nature, we have been unable to account for them within the OT framework the way we did for word-based hypocoristics. Their intricacy stems from the fact that they show an unpredictable gemination. To pave the way for a future possible analysis, we have reclassified them based on three criteria: first, the consonant that geminates; second, the consonants that surface; and third, those that do not. Though we have tried to capture few generalizations, we still believe that these hypocoristics need to be listed in the lexicon as they are.

Variation as Reflective of a Socio-pragmatic Function

Hypocoristics in MAm may be used in a young-to-old person communicative context, among equals, and in an adult-to-young person context. When used in a young-to-old person communicative context, which is most common, hypocoristics express respect toward the referent. A younger speaker should never address an older person by his/her full name but should use the appropriate hypocoristic form used for adults. Using the unsuitable form or full names to address adults violates the polite norm and may result in social rebuke or friction. The older person usually addresses a younger person by his full name or sometimes by a hypocoristic form for young people, when available. In this situation, hypocoristic forms denote smallness, even diminutiveness, of the referent. Consider the following examples illustrating this variation:

(28)	Full name	Young	Adult	
	ʕabdərˤrˤəhman	hani / bahhu / haman	hha, hhu	Abderrahman
	bəɫqasəm	baqqu	qqu	Belkacem
	ʕabdəlqadər	qada	qqa	Abdelkader
	ʕabdəslam	bassa / sassa	ssa	Abdesslam
	muħəmməd	ħəmmu	mmu	Mohamed
	buʕlam	ʕənnə	lla	Bouâlam
	xadiʒa	xiʒʒa	xxa / ʒʒa	Khadija
	fatˤima	ttima / bətti	tatta / titi	Fatima

Not all hypocoristic forms have two variants, that is, one for young people and another for old people. In fact, the great majority of hypocoristics are of the second type (i.e., forms for adults), suggesting that in MAm they are mainly used as status markers. Among equals, the standard mode used by adults to address each other is the use of hypocoristic forms. In contrast, young people address each other, mostly, by using their respective full names.

Age is the most important factor in determining the mode of address in MAm. We believe that it is the same in other languages or at least in English (Schneider 2003) and Akkan (Obeng 1997). In Akkan, the use of hypocoristic forms is governed by both age and socioeconomic status. Obeng (1997: 54) argues that their use

connotes friendliness, intimacy, similarity, solidarity, and the idea of belonging to a social group. Among equals, the use [...] is reciprocal. Where a difference of power or status or class association is involved, then the use of hypocoristic forms is nonreciprocal. A respectable superior can address a social inferior, if she/he chooses to, by his/her hypocoristic day-name. A social inferior, however, cannot address or refer to his/her superior by the superior's hypocoristic day-name unless it is prefaced by a deferential title.

In English, both age and social status are equally important in determining the mode of address (Ervin-Tripp 1972). However, the pragmatic function of the hypocoristic depends on the pattern of the truncation: simple truncated forms

vs. *y*-hypocoristics (Schneider 2003: 145). Schneider shows that *y*-hypocoristics “constitute the standard way” of addressing small children in English. However, among adults truncated names are the most common address forms and express the idea that the addresser perceives the addressee as his/her equal in terms of social status.

Finally, it should be noted that some hypocoristic forms in MAm became independent family names and are no longer used to refer to first names. Illustrative examples are given in (29) below:

(29)	Source name	Hypocoristic	
	saʕid	ʕəddɨ	Saïd
	bəlqasəm	qasu	Belkacem
	ʕabdəlɰafi	fafi	Abdelwafi
	ʕabdʕəlʔa	ʕəppu / ʕəlʔa	Abdellah
	ʕabdarʔahman	rʕəhhu	Abderrahmane
	ʕabdləʕziz	ʕəzzi	Abdelaziz
	ʕabdləʕlil	ʕəʕlil	Abdeljalil
	ʕabdəlqadər	qadi	Abdelkader
	ʕʕilali	lali	Jilali

Conclusion

The paper has surveyed hypocoristics in MAm with a view to first establishing an inventory of the different existing patterns and then attempting to account for the reduplicative and truncatory processes involved in their formation. Our objective has been to decide whether it is the root or the word that serves as a base for derivation.

Our study starts with a classificatory description of hypocoristic formation in MAm in section “[Hypocoristics in Moroccan Amazigh](#),” where patterns of hypocoristics have been established based on the nature of the base form and also on the behavior of their output forms. We have been able to identify five different patterns of hypocoristics: Patterns I and II, which resort to truncation of material from the left edge of the base, thus retaining the part which is right-anchored with the prosodic word. These forms consist of either two syllables or one syllable preceded by an initial geminate. For Patterns III and IV, we have identified forms that resort to the reduplication of one of the base segments of the root, mostly one or two consonants of a triconsonantal root. Pattern V hypocoristics are of a mixed type, showing variation between consonant reduplication and simple truncation to a disyllabic output without reduplication.

In terms of analysis, we have offered a preliminary account of the behavior of a number of hypocoristics in terms of the OT framework. The account was based on whether or not the element serving as a base to the derivation of hypocoristics is a root or a word (an output form).

We have shown that when the hypocoristic is word-based, the output form tends to be more predictable: the retained part is the right-hand element. Such a form is dictated by a number of constraints, the most important of which are constraint ANCHOR-R, requiring anchoring of the base and the truncated form, and the constraint FT-BIN, which forces reduplication of a root consonant of the input. We have been unable to show how the OT framework could account for hypocoristics that are root-based, because these show variation patterns, both in terms of consonant and vowel constituents. Instead, we have proposed to reclassify them based on the consonant that reduplicates and also on the root consonant that shows up in the output form.

Finally, we have pointed out to cases of variation, stating that these reflect a socio-pragmatic function regulating the relationship between adults and youngsters in a communicative context. We have shown that when youngsters address an adult, they can use hypocoristic names only but not the adult's full name. However, an adult addresses a youngster using this latter's full name, or a hypocoristic for youngsters, if it exists.

Appendix

Hypocoristics in MAm

The data below are representative examples of the different hypocoristic patterns in MAm.

Pattern I: Right-anchored CVCV hypocoristics

Name	Hypocoristics	
r ^h abiʃa	biʃa	Rabiâa
məʃzuza	zuza	Maâzouza
lmadani	dani	El Madani
məns ^h ur ^h a	sula	Mansoura
kəzzaza	zaza	Kezzaza
ʃabdəlkarim	kari	Abdelkarim
buʃəzza	buza	Bouazza
ʃabdəlɰafi	ʃafi	Abdelɰafi
lʃar ^h abi	ʃabi	El Arabi
lyalja	ɣaja	Ghalia
tahhat	taha	Tahhat

Pattern II: Right-anchored C_iC_iV hypocoristics with initial gemination

Name	Hypocoristics	
boufəɭʒa/xadiʒa	ʒʒa	Boufelja/Khadija
ʃi(j)fa	ʃʃa	Aïcha
məs ^h t ^h afa	f ^h f ^h a	Mustafa

(continued)

Name	Hypocoristics	
xira	xxa	Kheyra
ʃixa	xxa	Chikha
ʃaʃur ^ʕ	ʃʃu	Achour
s ^ʕ əd ^ʕ d ^ʕ iq	dɖi	Saddiq
ʃabdəlqadər	qqa	Abdelkader
ʃabdər ^ʕ əhman	hha	Abderrahmane
ʃabdəsslam	ssa	Abdelssalam
ʒəmʃa	ʃʃa	Jomâa
kəzzaza	zza	Kezzaza
buʃlam	lla	Bouâlam

Pattern III: CVC_iC_iV hypocoristics with medial gemination

Name	Hypocoristics	
mbar ^ʕ ka	bakka	M'barka
zzəhr ^ʕ a	zazza	Zahra
buzijjan	bazza	Bouziyane
ʒəmʃa	ʒaʃʃa	Jomâa
ʃabdəlʃʃah	baʃʃa	Abdellah
ʃabdəsslam	sassa	Abdelssalam
ʃabdəsslam	bassa	Abdelssalam
xatima	xamma	Khatima
xatima	mamma	Khatima
fat ^ʕ ima	tatta	Fatima
rqiija	jaɣja	Rqiya
hməd	dadɖa	Ahmed

Pattern IV: CəC_iC_iV hypocoristics with medial gemination

Name	Hypocoristics	
mhamməd	həmma	M'hamed
bəxta	xəttə	Bekhtha
ʃabdəlʃʃah	bəlla	Abdellah
tbuʒʒat	bəʒʒa	Bahija
jamina	mənna	Yamina
buʃlam	ʃənna	Bouâlame
ʃumar ^ʕ	ʃəm ^ʕ m ^ʕ a	Omar
bəlxir	bəxxi	Belkhir
bəlʃid	bəʃʃi	Belaïd
hasan / husin	həssi	Hassan/Houcine
ʃar ^ʕ bi	ʃəvvi	El Arbi
fat ^ʕ ima	bətti	Fatima
saʃid	ʃəɖɖi	Saïd

Pattern V: Mixed CVCV hypocoristics

Name	Hypocoristic	
sliman	sima	Slimane
ʕabdləmʒid	ʒida	Abdelmajid
hməd	hida	Ahmed
muħəmməd	muda	Mohamed
məbrʕuka	buka	Mabrouka
brʕahim	bahu	Brahim
bəlqasəm	qasu	Belkacem
mulud	lulu	Mouloud
tʔʔjəb	bubu	Taïb
muħammad	madi	Mohamed
ʕabdəlwf	fafi	Abdelwafi
ʕabdərʕəħman	hani	Abderrahmane
tʔʔjəb	bibi	Taïb

References

- Abu-Mansur MH (2019) The role of the OCP and syllable structure in Makkan Arabic hypocoristics. *Ling Lit Stud* 7(1):31–38
- Al Ghadi A (1990/2014) Moroccan Arabic plurals and the organization of the lexicon. Publications of the Faculty of Letters, Rabat
- Alber B (2010) An exploration of truncation in Italian. *Rutgers Work Pap Ling* 3:1–30
- Alber B, Arndt-Lappe S (2012) Templatic and subtractive truncation. In: Trommer J (ed) *The morphology and phonology of exponence*. Oxford University Press, Oxford, pp 289–325
- Álvarez JS (2015) The phonology and morphology of Spanish hypocoristics. MA thesis, The Arctic University of Norway
- Azizi S (2013) Le nom de personne dans l'oasis de Figuig: Un système de codification des relations sociales. In: Belghazi H (ed) *La culture amazighe: Réalités et perceptions*. Actes du colloque organisé par le CEAS-IRCAM et la FLSH de Meknès. Publications de l'IRCAM, Rabat, pp 41–62
- Bat-El O (1994) Stem modification and cluster transfer in modern Hebrew. *Nat Lang Ling Theory* 12:571–596
- Ben-Abbas M (2003) Variation et emprunts lexicaux: Etude sociolinguistique sur le parler amazigh de Figuig. Unpublished doctoral thesis, Sidi Mohamed Ben Abdellah University, Fes
- Benamara H (2011) Contes berbères de Figuig (Sud-est marocain). *Berber studies* 34. Rüdiger Köppe Verlag, Köln
- Benamara H (2013) Dictionnaire amazighe-français: Paler de Figuig et ses régions. Publications de l'IRCAM, Rabat
- Benmamoun E (1999) Arabic morphology: the central role of the imperfective. *Lingua* 108: 175–120
- Bensoukas K (1994) Tashlhit Agentive nouns in Tashlhit Berber: an optimality-theoretic account. Unpublished D.E.S. Thesis, Mohammed V University, Rabat
- Bensoukas K (2018) Against the consonantal root in Tashlhit. *Asinag* 13:53–90
- Benua (1995) Identity effects in morphological truncation. In: Beckman J, Urbanczyk S, Walsh L (eds) *University of Massachusetts occasional papers in linguistics* 18: papers in optimality

- theory. Graduate Linguistic Student Association, Amherst, pp 77–136. <http://roa.rutgers.edu>, ROA-74
- Boudlal A (2001) Constraint interaction in the phonology and morphology of Casablanca Moroccan Arabic. Unpublished Doctorat d'État dissertation, Mohammed V University, Rabat. <http://roa.rutgers.edu>, ROA#650
- Boudlal A (2018) Root-and-pattern morphology revisited: verb stem bimoraicity and stem-based morphology in Moroccan Arabic. *Asinag* 13:165–190
- Boudlal A, Yeou M (2019) A survey of hypocoristics in Moroccan Amazigh. In: Erguig R, Mokhtari N, Yeou M (eds) *Cultures and languages in contact V*. Editions Bouregreg, Rabat
- Davis S, Zawaydeh BA (1999) A descriptive analysis of hypocoristics in colloquial Arabic. *Lang Ling* 3:83–98
- Ervin-Tripp S (1972) Sociolinguistics. In: Pride JB, Holmes J (eds) *Sociolinguistics: selected readings*. Penguin, Harmondsworth, pp 223–240
- Heath J (1987) Ablaut and ambiguity: phonology of a Moroccan Arabic dialect. State University of New York Press, Albany
- Kossmann MG (1997) *Grammaire du berbère de Figuig (Maroc Oriental)*. Peeters, Louvain/Paris
- McCarthy J (1981) A prosodic theory of nonconcatenative morphology. *Ling Inq* 12:373–418
- McCarthy J, Prince A (1993) Prosodic morphology I: constraint interaction and satisfaction. Ms. University of Massachusetts, Amherst and Rutgers University. <http://roa.rutgers.edu>, ROA# 482
- Newman P, Ahmad M (1992) Hypocoristic names in Hausa. *Anthropol Ling* 34(1–4):159–172
- Obeng SG (1997) From morphophonology to sociolinguistics: the case of Akkan hypocoristic day-names. *Multilingua* 16(1):39–56
- Oussikoum B (2013) *Dictionnaire amazighe–français: Le parler des Ait Wirra – Moyen Atlas*. Publications de l'IRCAM, Rabat
- Prince A, Smolensky P (1993/2004) *Optimality theory: constraint interaction in generative grammar*. Blackwell, Malden/Oxford, UK. (Revision of 1993 technical report, Rutgers University Center for Cognitive Science. <http://roa.rutgers.edu>, ROA#537)
- Saa F (2010) *Quelques aspects de la morphologie et de la phonologie d'un parler amazighe de Figuig*. Etudes n°12. Publications de l'IRCAM, Rabat
- Schluter KT (2013) *Hearing words without structure: subliminal speech priming and the organization of the Moroccan Arabic lexicon*. Unpublished Doctoral dissertation, University of Arizona
- Schneider K-P (2003) *Diminutives in English*. Niemeyer, Tübingen
- Serhoual M (2002) *Dictionnaire tarifit-français et essai de lexicologie amazighe*. Unpublished doctoral thesis, Université Abdelmalek Essaâdi, Tetuan
- Taïfi M (1992) *Dictionnaire tamazight-français: Parlers du Maroc central*. l'Harmattan- Awal, Paris
- Taïfi M (2016) *Le dictionnaire raisonné berbère-français: Parlers du Maroc*. Publications de l'IRCAM, Rabat
- Ussishkin A (1999) The inadequacy of the consonantal root: modern Hebrew denominal verbs and output-output correspondence. *Phonology* 16:401–442
- Yeou M (2022) *Dictionnaire Amazighe-Français. Parler de Figuig*. Publications de l'IRCAM, Rabat
- Yeou M, Honda K, Maeda S (2011) A palatographic, photoglottographic and airflow study of some coronal geminates in Figuig Berber. In: Lee W-S, Zee E (eds) *Proceedings of the XVII international congress of the phonetic sciences*, pp 2216–2219, Hong Kong, China

Comparaison et reconstruction en berbère 32

Exemple du lexique géographique

Abdallah El Mountassir

Contents

Structuration et organisation du lexique	683
Quelques données berbères	684
Comparer pour reconstruire	686
Exemple de la base ȲR et la notion de “dureté, sécheresse”	688
Lieu sec et pierreux, endroit dur, terrain rocailleux et aride,	689
<i>azayar, tazayart, pl. izuyar</i>	689
<i>akerkur, takeṛkur</i>	692
Sécheresse, plage et aire à battre les céréales	693
Lexique géographique berbère: une diversité de désignations et une unité de signification ...	694
Références bibliographiques	696

Structuration et organisation du lexique

Dans la description des langues, les linguistes (structuralistes) s'accordent à admettre que le lexique d'une langue n'est pas un simple répertoire des mots isolés les uns des autres. Comme les phonèmes (en phonologie) et les formes grammaticales (en morphologie et en syntaxe) sont structurables, le lexique peut l'être aussi. Le linguiste George Mounin (1975: 49–50) affirme dans ce sens que « [...] tous les linguistes d'aujourd'hui demeurent d'accord *qu'il doit y avoir une organisation quelconque* du lexique, et de son contenu: les significations. Tous demeurent d'accord qu'il est impossible de penser que les mots sont présents d'une manière ou de l'autre dans notre tête sous la forme d'éléments totalement isolés les uns des autres. C'est cette conviction qui s'exprime quand on répète qu'une langue n'est pas

A. El Mountassir (✉)
Université Ibn Zohr, Agadir, Morocco
e-mail: a.elmountasser@uiz.ac.ma; elmountassir@hotmail.com

une nomenclature (Saussure), un répertoire, un inventaire (Martinet), un sac à mots (Harris), un empilement de noms (Whorf). »

La méthode d'analyse qui est le plus souvent suivie dans la structuration d'un lexique quelconque est basée sur le principe de l'opposition entre les structures formelles et les structures sémantiques: des différences de forme des unités linguistiques ne sont pertinentes que si elles entraînent des différences au niveau du sens. Pour les structuralistes, il est donc possible de structurer le lexique en délimitant des champs restreints de significations ou champs sémantiques. Un champ sémantique est un ensemble de termes qui se rapporte à un même domaine d'expérience. Il importe de préciser ici que chaque champ sémantique est intimement lié au contexte socioculturel. En d'autres termes, un champ sémantique constitue une catégorisation et une classification d'un domaine d'expérience propre à une communauté. Un champ sémantique d'habitat, par exemple, ne peut être conçu d'une manière identique dans toutes les langues ou dans les communautés culturelles différentes.

Dans cette perspective, la structuration du lexique, qui varie d'une langue à l'autre, est donc un véritable dévoilement de la pensée. Les linguistes, et plus particulièrement les ethnolinguistes, s'accordent à reconnaître dans l'étude du lexique un meilleur moyen de pénétrer dans l'univers mental des locuteurs: les mots finissent toujours par reconstituer les systèmes de pensée et par décrypter le monde. Comme le souligne Jacqueline Picoche (1993: 39) « le lexique n'est pas un simple inventaire, dûment étiqueté, des réalités extra-linguistiques, mais une interprétation de cette réalité. »

Dans leurs analyses du lexique d'une société donnée, les chercheurs ont toujours été sensibles aux champs morphosémantiques. Il s'agit d'identifier et d'étudier des rapports formels et sémantiques dans une série de mots. Rappelons qu'un champ morphosémantique implique l'existence d'un ensemble de groupes de mots qui présentent des éléments formels (signifiants) et sémiques (signifiés) communs. Le concept de champ morphosémantique dans les langues suppose ainsi la convergence des deux structures morphologique et sémantique; ce qui signifie que les mêmes catégories sémantiques sont exprimées par les mêmes catégories morphologiques (voir ci-dessous l'étude de la base ΥR). L'ensemble d'une série de mots d'un champ morphosémantique remonte historiquement à une base commune qu'on désigne par l'étymon. Cette base constitue ainsi une matrice et un support de famille des mots (Haudry 1985).

Quelques données berbères

Dans les études sémitiques, David Cohen (1975) avait attiré l'attention sur l'existence de plusieurs séries de racines bi-consonantiques qui évoquent telle ou telle signification fondamentale. A partir par exemple d'une base radicale BT qui véhicule la notion centrale de « couper, fracturer », D. Cohen a établi plusieurs groupes de racines BTT, BTR, BTK, BTL, ... Toutes ces racines sont associées au champ sémantique de « couper, fractionner, diviser, fragmenter, ... ». Nous avons

donc ici un groupe de racines tri-consonantiques qui ont en commun la base BT, considérée comme le noyau sémique. En d'autres termes, nous avons les mêmes structures formelles qui véhiculent, dans divers contextes, les mêmes valeurs sémantiques.

Dans le domaine berbère, des travaux sur la structuration du lexique et l'élaboration des champs morphosémantiques ont été publiés sur ce sujet. (Nous citons ici Galand-Pernet (1963–1966, 1984); Chaker (1979, 1995); El Mountassir (1996, 2002); Galand (1997); et Naït-Zerrad 2000.) Dans ses travaux sur les bases radicales en berbère, Salem Chaker a mis en évidence une liste d'une trentaine de bases bi-consonantiques. Chaque base véhicule une notion centrale: FR « notion de séparation » (*frs* « nettoyer, émonder, ... », *frn* « trier », *fruri* « éclore, sortir de son enveloppe », ...), GM « notion de (re)cueillir, rassembler » (*gmr* « chasser, cueillir, récolter, ... »), KR « notion de rondeur, enroulement » (*krs* « nouer », *krd* « emballer », *krf* « être entravé »), NY « notion de destruction » (*ny* « tuer », *nyd* « écraser »), etc. Avec l'établissement de ces séries de racines radicales, nous ne pouvons que constater les rapports existants entre la forme et le sens des unités lexicales.

Concernant le lexique berbère de l'espace, nous avons étudié dans nos travaux antérieurs (El Mountassir 1998, 2001, 2002, 2005a, b) quelques bases radicales à ce sujet. Chaque base bi-consonantique génère plusieurs groupes de racines qui sont associées à un champ sémantique particulier dans le domaine spatial et géographique. A partir de la base LY, par exemple, nous avons une série de quatre racines LK, LQ, LY et LX qui évoquent les diverses nuances sémantiques de « terre alluviale, boue, humidité, terrain marécageux, vaseux, ... »:

lley « lécher », *taluyi* « mollesse », *ilxix* « être imbibé d'eau (terre) », *allay*, pl. *allayn* « fond (d'un trou, d'un puits, ...) », *iliy*, pl. *iliyn* « petite rivière » (Le terme *iliy* est attesté en toponymie dans le sud-ouest marocain. C'est le cas de la localité d'Iliy dans le versant ouest de l'Anti-Atlas, située à une soixantaine de kilomètres au sud-est de Tiznit. Rappelons que cette localité est célèbre par son château *Tigmmi n-Iliy* qui a joué un rôle important dans le commerce transsaharien durant la période XVII–XIX^e siècle.) (tachelhit, Maroc central)

allay, *tallayt* « boue, terre boueuse », *tilkki* « terrain marécageux », *amllaku* « terre alluviale » (tachelhit) (La toponymie de l'Anti-Atlas occidental offre un grand nombre d'exemples avec le terme *amllaku*, ou la forme du féminin *tamalukt/tamllakut*. Nous citons ici l'exemple de *Tamalukt* qui est le nom d'un petit village au nord de l'oued *Ulyas* à *Id gg° Arsmouk*.)

alxix, pl. *ilxixn* « sables mouvants, terrain humide et boueux » (Maroc central)
ilqiq « être mou », *alyhuy* « mou et humide », *luy* « être trouble, boueux », *allay* « fond de vase, de boue », *tallayt* « boue de terre glaise » (kabyle, Dallet 1982: 458)

talayyt « vase gluante » (kabyle, Dallet 1982: 459)

talak « argile, terre grasse », *alaku*, pl. *alakutn* « vase, boue qui se dépose au fond des eaux », *ilouk* « sables mouvants » (touareg de l'Ahaggar, Foucauld 1951: 1029–1030), etc.

Tous ces mots, dispersés dans des aires géographiques différentes, sont rattachés au niveau diachronique à la base LY (latérale + vélaire) qui véhicule la notion centrale de « humidité, mouillage ». Nous avons ici une matrice sémantique fondamentale et productive en berbère. (A propos de cette base LY, voir El

Mountassir (2002).) Il est important de rappeler que toutes ces racines, dont la latérale *L* constitue l'élément constant, sont apparentées et que les consonnes *y*, *q*, *k* et *x* sont considérées comme des variantes dans le système phonologique diachronique du berbère (Chaker 1995).

Par ailleurs, dans notre étude sur le lexique de l'espace en berbère, nous avons mis en évidence que ce vocabulaire présente beaucoup de mots isolés qui ont rompu, par suite aux évolutions phoniques et sémantiques, des liens avec leurs familles. Notre objectif dans ces travaux est d'essayer de regrouper ces termes isolés dans des familles de mots en identifiant les liens formels et sémantiques entre eux. C'est le cas par exemple de la série des mots comme *ifri* « grotte, cuvette (terrain), gîte », « aiguille rocheuse très mince » (touareg de l'Ahaggar), *tafrawt* « dépression de terrain, cuve, bassin de réception d'un puits », *iferd* « cuve, étang, bassin creusé destiné à retenir l'eau de la pluie », *fruri* « s'émietter, s'effriter », *fri* « déchirer » (kabyle), *frem* « être ébréché, édenté », *frekek*, *frekes* « se craqueler, se fissurer », etc., qui présentent bien des relations de forme et de sens. Toutes ces unités lexicales, ayant en commun la séquence consonantique FR, appartiennent à la même famille lexicale et sémantique qui évoque, dans le domaine géographique, des espaces et des terrains présentant des dépressions et accidents topographiques (cluse, cuve, cuvette, grotte, ...). En d'autres termes, ce groupe de mots forme un champ morphosémantique cohérent.

Comparer pour reconstruire

A l'intérieur d'un champ morphosémantique, les unités lexicales ne sont analysables, dans la plupart des cas, qu'en diachronie. D'où la nécessité et l'importance de la comparaison interdialectale. Pour le traitement du lexique d'une langue, la comparaison est un outil fondamental. Une étude des champs lexicaux ou sémantiques en berbère ne peut apporter des résultats satisfaisants si elle est limitée à un seul parler. La méthode consiste à établir un inventaire de mots, aussi complet que possible, attestés dans des parlers différents. Dans le domaine du vocabulaire de l'espace en berbère par exemple, une telle démarche comparative permet de mettre en évidence des modes de formation lexicale et d'appréhender des changements et l'évolution des termes sur le plan phonique (usure phonétique) et sémantique.

Notre analyse du vocabulaire de l'espace est donc basée sur le principe de la comparaison des unités linguistiques dispersées dans des aires géographiques diverses. Pour pouvoir reconnaître la parenté entre ces différentes unités, la comparaison interdialectale s'impose, car sur le plan synchronique, les mots paraissent toujours indépendants les uns des autres. Dans ce contexte, on ne peut parler de méthode comparative sans évoquer la reconstruction. Les deux démarches sont liées. A ce propos, Galand-Pernet (1985: 15) précise: « [...] Quant au 'berbère' au sens de protolangue, ayant existé (sous quelle forme ?) à une époque préhistorique, on peut affirmer qu'à partir d'un noyau commun il s'est différencié au cours de la protohistoire de façon à produire des variétés linguistiques qui ont suivi leur propre évolution, sur des territoires aujourd'hui dispersés de l'Égypte au Maroc. On se sert

donc de la comparaison pour reconstruire les traits anciens de ce proto-berbère, comme on le fait pour l'indo-européen. »

La comparaison est également indissociable à l'explication des unités lexicales. Dans le champ des études berbères, la comparaison interdialectale rend des grands services aux linguistes dans la mesure où cette comparaison permet d'expliquer certaines évolutions phonétiques, morphologiques et sémantiques à l'intérieur d'un parler particulier. Sur le plan sémantique par exemple, l'étude de certains mots est assez intéressante pour discerner le degré de leur évolution du sens. Prenons le cas du terme *tigmmi* qui a le sens de « maison » en tachelhit. La comparaison avec d'autres parlers berbères, où le même terme est attesté, permet de comprendre comment le sens actuel de « maison », en tant que “construction”, en tachelhit est lié à une signification antérieure en rapport avec la notion d'« unité sociale »:

- « domaine, maison, tente », Ayt Hdidou (Azdoud 1985: 66)
- « douar, groupe de tentes dressées et disposées en cercles; groupe de gens placés en rond; place centrale d'un village ou d'un douar », Maroc central (Taïfi 1991: 157)
- « maison-forte, châtelet », vallée de Dadès
- « village, campement, bivouac », Ayt Mgild
- « vaste terrain de culture », kabyle (Dallet 1982: 260)
- « centre de douar où sont parqués les animaux », Zemmour
- « jardin irrigué », Mzab (Delheure 1984: 69).

Il ressort de ces différents sens « maison, tente, campement, jardin, village » que le terme *tigmmi* renferme l'idée de « espace approprié par l'homme pour ses besoins domestiques ». La signification initiale de *tigmmi* ne peut s'interpréter que dans le cadre d'une structure sociale berbère bien déterminée. Cette signification ne traduit, en aucun cas, le concept matériel de maison (en tant que maison construite et bâtie). La conception matérielle de *tigmmi* est le résultat d'une évolution sociale de la société berbère. (Pour plus de détail à ce sujet, voir El Mountassir (2010).)

Le sens de *tigmmi* en tant que “construction” est donc bien postérieur à *tigmmi* en tant que “unité sociale (groupe familial)”. Mais, l'évolution de la langue en parallèle avec l'évolution sociale a fait de sorte que ce terme désigne aujourd'hui maison-bâtie en tachelhit. Ce type de changement sémantique révèle le phénomène qui s'est produit dans d'autres langues en dehors du monde berbère. Les études de l'évolution des langues indo-européennes par exemple ont fini par distinguer le groupe de résidence domestique et l'espace qu'il occupe. La confusion entre les deux concepts -social et spatial- avait entraîné des difficultés dans les travaux des étymologistes dans le domaine du vocabulaire de l'habitation. Comme l'a bien souligné Emile Benveniste (1955: 21): « les unités sociales ont été identifiées au cadre spatial de leur habitat. Ces exemples montrent comment des dénominations sociales ont été transférées à des groupements territoriaux. Ainsi, la notion de “famille” et celle de “habitation” (bâtie) se sont rapprochées puis associées (...) il s'est produit des contaminations qui obscurcissent et rend souvent indécise la répartition étymologique » (Meillet 1925).

La remarque de Benveniste est tout à fait valable pour le berbère. Nous verrons plus loin dans d'autres exemples comment les concepts sociaux ont été transférés aux concepts spatiaux et géographiques.

Exemple de la base γR et la notion de “dureté, sécheresse”

Le lexique géographique berbère présente un nombre considérable de termes désignant les variétés des sols, les formes du relief et les types de terrains. Nous retrouvons dans ce lexique un vocabulaire commun à la plupart des parlers berbères qui sont dispersés et éloignés géographiquement. Nous avons là une richesse lexicale qui n'a pas encore bénéficié d'études rigoureuses et systématiques de la part des chercheurs berbérissants. L'étude de ce lexique, qui contient un vocabulaire ancien, est un meilleur moyen de restituer des fonds lexicaux et sémantiques communs à l'ensemble des parlers berbères.

Dans une bonne partie des parlers berbères, le vocabulaire abondant qui désigne un espace dur, aride et sec (lieu sec, terrain pierreux, bord de mer, plaine sèche, ...) est construit à partir de la base γR (base bi-consonantique) qui véhicule la notion de « dureté, sécheresse ». Nous avons ici une matrice de base fondamentale et productive en berbère.

A cette notion se rattache un grand nombre de termes formant un champ lexical immense. Cet inventaire lexical est constitué d'une centaine de mots appartenant à des aires dialectales diverses. L'ensemble de ces termes forme un champ morphosémantique cohérent. Rappelons qu'un champ morphosémantique est constitué d'un ensemble de mots qui sont apparentés par leurs formes et par leurs sens. Chaque champ morphosémantique implique l'existence d'un ensemble de mots qui présentent des éléments formels et sémiques communs. Ajoutons que l'ensemble de la série des mots d'un champ morphosémantique remonte au niveau diachronique à une base commune (qu'on désigne par l'étymon). Celle-ci constitue le niveau le plus profond d'une famille de mots.

Dans ce qui suit, nous tenterons de reconstruire ce champ morphosémantique qui exprime la notion de « dureté, sécheresse », en examinant quelques relations morphologiques et sémantiques entre, d'une part, les divers termes de ce champ et, d'autre part, entre l'étymon et les mots qui en dérivent. Le but de cette analyse est d'appréhender les différents types de relations de parenté qu'entretiennent ces termes. S'il est plus facile de déceler par exemple la relation morphologique et sémantique entre le verbe *yar* « être sec, dur » et ses dérivés: *syer* « sécher, durcir », *asyer* « fait de sécher, dureté », cette relation est moins facile à établir entre les mots *tayart* « plage », *akeṛkur* « endroit, terrain pierreux », *azayar* « plaine, extérieur, dehors », pl. *izuyar* « désert » (Mزاب, Ouargla), *azqqr* « bûche de bois », (kabyle), ... etc. La relation qui unit ces unités lexicales est strictement diachronique ce qui n'est pas, normalement, perceptible pour les locuteurs. C'est en établissant une notion commune de ces mots et en les regroupant dans un même champ morphosémantique qu'on peut alors se rendre compte de leur parenté et leur interdépendance formelle et sémantique.

Par ailleurs, comment comprendre par exemple le sens étymologique de *tayart* « plage » en tachelhit sans tenir compte de différents sens que ce même terme possède dans d'autres parlers berbères ? La solution est de collecter la totalité d'acceptions sémantiques de ce terme attesté dans divers parlers pour se rendre compte de l'existence d'un noyau sémique. C'est également par l'étude comparative qu'on

peut repérer l'évolution sémantique et formelle d'un terme et par là, son « épaisseur » diachronique. Nous adoptons donc ici une approche qui dépasse le cadre figé de l'analyse synchronique pour tenter une analyse historique et diachronique par le biais de la comparaison interdialectale. Cette approche permet de restituer et de comprendre les permanences et les changements des formes linguistiques.

Ajoutons que ce champ morphosémantique est formé de racines bilitères *ȚR*, *QR*, *KR*, *GR*. Toutes ces racines sont apparentées. Les consonnes *Ț*, *Q*, *K* et *G* sont considérées ici comme des variantes dans le système phonologique diachronique du berbère (Chaker 1995).

Lieu sec et pierreux, endroit dur, terrain rocailleux et aride,...

akerȥur, fém. *takerȥurt* « endroit sec et pierreux, tas de pierres », tachelhit
takerȥurt « sol pierreux, dur », tachelhit (plaine d'Achtoukn)

takerakart « terre sèche », tachelhit (Laoust 1920: 262)

akerkur, pl. *ikerkurn* « tas de pierres »; *takerkurt*, *amquqr* « petit tas de pierres », Maroc central (Taïfi 1991: 344)

Takerkurt n-Ahaggar « massif montagneux central du pays des Kel-Ahaggar », (Foucauld 1951: 875)

aqerqar « endroit sec et pierreux », kabyle (Dallet 1982: 670)

teyart « fait d'être sec, de se sécher », par ext. « endroit sec », touareg de l'Ahaggar (Foucauld 1951: 1752)

atyer « endroit plat à sol argileux », touareg de l'Ahaggar (Foucauld 1951: 1754)

teyeryert « aire pour battre le grain », touareg de l'Ahaggar (Foucauld 1951: 1755)

tiyeryert « sol de maison », kabyle (Dallet 1982: 623)

ayerȥur « élargissement de vallée, à sol uni et dur, en plaine », touareg de l'Ahaggar (Foucauld 1951: 1755)

azȥar « plaine; plaine sèche », *tazȥart* « petite plaine », kabyle (Dallet 1982: 952)

azȥar, pl. *izuyarn* « plaine; plateau. Grande plaine servant de pâturage au bétail pendant la transhumance d'hivers », Maroc central (Taïfi 1991: 799)

azȥar « plaine », tachelhit

azȥar « extérieur, dehors », pl. *izuyar* « désert », Mzab (Delheure 1984: 253)

azȥar, pl. *izuyar* « extérieur, dehors », par ext. « désert, hors de l'oasis », ouargli (Delheure 1987: 395)

tayȥart, pl. *tiȥaryaren*, « endroit de terre séché et durci par le soleil », touareg de Niger (Prasse et al. 1998: 121)

azȥar, *tazȥart*, pl. *izuyar*

Dans cet inventaire, on remarque que le terme *azȥar*, fém. *tazȥart* est attesté dans plusieurs parlers berbères avec des sens différents. Dans certains parlers (tachelhit, Maroc central, kabyle), on traduit souvent *azȥar* par « plaine, plateau » qui

s'oppose à *adrar* « montagne »; et dans d'autres (Mزاب, Ouargli) le même terme désigne « dehors, extérieur, désert ». La comparaison interdialectale et l'analyse sémantique montrent que le noyau sémique du terme *azayar* ne contient aucune notion de « plat, platitude » ou de « dehors ». Quelle est donc l'idée initiale que véhicule le sémantisme d'*azayar* ? Pourquoi ce terme désigne-t-il ici « plaine, plateau » et là « dehors, extérieur » ?

Par ailleurs, on relève de nombreux toponymes avec le terme *azayar*, ou sa forme du pl. *izuyar* dans le sud-ouest du Maroc et dans les zones présahariennes (les oasis de *Tata* et d'*Aqqa*). La plupart de ces lieux étaient des anciens pâturages pour les habitants de l'Anti-Atlas oriental. Emile Laoust (1940: 33) a mentionné le nom *Tazayart*, forme féminine d'*azayar*, dans le Haut Atlas central. C'est « le nom d'un plateau qui se dresse au sommet de murailles puissantes de Wankrim. Ce plateau, *nu, vide*, compris entre 3800 et 3980m est un *désert de pierres* ». La description de Laoust à propos du toponyme *Tazayart* résume bien le sémantisme de ce terme. En effet, cette définition met en évidence les traits sémantiques de « sécheresse, dureté, nudité, aridité ».

A travers les données linguistiques présentées dans notre inventaire, nous comprenons que le mot *azayar* est une forme nominale en *as-* construite à partir du verbe *syer* « sécher, faire sécher, durcir ». Ce verbe est lui-même une forme dérivée en *s-*, à valeur causative, à partir de *yar* «être sec, dur». Le même verbe est connu dans plusieurs parlers:

- yar, tɣar* (inac.) « être dur, sec (objet, lieu, ...) », *syer* « sécher, durcir », tachelhit
qqar « être sec », *syer* « durcir, rendre dur, faire sécher », kabyle (Dallet 1982: 622)
qerqer « être nu, désert (terrain) », kabyle (Dallet 1982: 670)
yar « être sec, se sécher », par ext. « être durci, se durcir, être dur (sol, terrain, pain, bois, ...) », *yeryer* « rendre sec et dur », *ayeryer* « fait de rendre sec et dur », *syer* « sécher, faire sécher », *asyer* « fait de rendre sec », touareg de l'Ahaggar (Foucauld 1951: 1751, 1753–1754)
yar « être sec, se sécher; être dur, durcir », *seyer* « sécher », *aseyer* « fait de sécher », touareg de Niger (Prasse et al. 1998: 118)
qqur « être sec », Siwa (Laoust 1932: 294)
qqar « être sec, dessécher, faner, être aride (terre) », *syer* « sécher, rendre sec, faire sécher », *aqquurar* « sec, aride, dur » (adj.), Maroc central (Taïfi 1991: 198)
qqar, tɣar, « être sec, être dur, durcir, être raide », *syr* « sécher, durcir », *ttusyr* « avoir été séché, durci », *ayurar*, pl. *iyurar* « le dur, le sec (adj.) », Ayt Hdiddou (Azdoud 1985: 181)
qqar « être sec, sécher, être dur, durcir », *sseqqar* « faire sécher, faire durcir », Mزاب (Delheure 1984: 162)
qar, ɣara « être sec », *syer, syara* « faire sécher », Beni-Snous (Destaing 1914: 321)
eyru « dessécher légèrement à la surface », *ɣruyru* « être presque à sec (lieu, puits, source, ...) », *yeregget* « être à sec, ne plus avoir d'eau (cours d'eau, ruisseau, source, puits) », *seqqergget* « mettre à sec », touareg de l'Ahaggar (Foucauld 1951: 1756–1757)
qqar « être sec, sécher, être dur, durcir », *ssqar* « faire sécher », ouargli (Delheure 1987: 242)
yeryer « être aplani, damé, aménagé en plate-forme, en terrasse », *ssyreyr* « aplanir, aménager en terrasse », *tɣyryin* « aire aplanie, plate-forme », ouargli (Delheure 1987: 242)
teyart « fait d'être sec, de se sécher », touareg de l'Ahaggar (Foucauld 1951: 1752)
teyiret « être desséché et durci (sol argileux) », touareg de l'Ahaggar (Foucauld 1951: 1754)

yyeryer « aménager une plate-forme », kabyle (Dallet 1982: 622)
aquran, fém. *taqurant* « sec, dur(adj.) », *iquranen* « période de l'été où tout est sec », kabyle (Dallet 1982: 622).

Pour les populations de la montagne, le terme *azayar* fait partie du vocabulaire spatial spécifique à leur mode de vie. Ainsi, la vallée du Souss, que les communautés de l'Anti-Atlas occidental *Ibudrarn* découvrent de leur montagne, est appelée *azayar*. Ce même terme désigne également pour les populations du Haut-Atlas central la grande plaine ensoleillée de Marrakech, connue aujourd'hui sous le nom de Lhaouz. La grande vallée de Sebou, au nord du Moyen-Atlas, portait jadis le nom *Azayar*. Cet ancien toponyme est remplacé aujourd'hui par le terme arabe *Lyarb* (litt. ouest). Pour toutes ces populations de montagnes, *azaġar* désigne un territoire (plaine, plateau, vallée) où elles mènent leurs troupeaux pendant les périodes de l'hiver. *Azayar* constitue ainsi pour elles un refuge contre le froid. Ce terme est donc en rapport au mode de vie et à l'expérience propres aux transhumants.

Chez les Ayt Hdidou, populations du Haut-Atlas oriental, *azayar* est un terme qui fait partie du lexique de l'orientation géographique (Azdoud 1985: 262). *Azayar* désigne ainsi dans ce milieu la direction ouest. Or, il est important de préciser que *azayar* est d'abord le nom du territoire situé à l'ouest du pays Ayt Hdidou. Autrement dit, la direction ouest est nommée ici à partir du nom du territoire. Nous avons ici un principe de la dénomination des directions géographiques qui est connu dans d'autres sociétés berbères. (Le vocabulaire de l'orientation en berbère est fondé essentiellement sur le vécu et la réalité géographique propres à chaque milieu. Ainsi, si le terme *tagut* « brume, vent frais de l'ouest » désigne aussi la direction ouest en tachelhit, c'est parce que, dans la plaine de Souss, la brume vient de l'ouest qui est la direction de la mer. Pour d'autres localités du Haut-Atlas occidental, la direction ouest est désignée par *lbhr* « la mer ». En ce qui concerne d'autres sociétés berbères, le terme *ataram* par exemple désigne chez les touaregs du Niger la direction ouest; or ce même terme signifie d'abord « le bas, l'aval » car c'est vers l'ouest que le soleil « tombe ». Chez les mêmes populations, la direction de l'est est désignée par *enneg*, terme qui signifie à l'origine « le haut ». A propos du terme *ataram*, mentionnons ici l'appellation *Kel Ataram* (= gens de l'ouest) qui désigne les touaregs *Kel Tamashaq* du Niger. (V. P. Galand-Pernet 1992; D. Casajus 1981; E. Bernus 1982.))

Les données linguistiques présentées plus haut montrent que, dans certains parlers berbères, *azayar* est associé à la notion de « lieu sec, inculte » qui se traduit dans la langue par les traits sémiques de « espace desséché, aridité, ailleurs, extérieur, espace non humain ». Il s'agit souvent des populations des oasis pour qui *azayar*, pl. *izuyar*, désigne le territoire situé à l'extérieur de l'oasis et à tout espace d'habitat humain. D'où le sens de « dehors, extérieur » et par extension « désert », comme c'est le cas des communautés des oasis de Mzab (Delheure 1984: 253) et de Ouargli (Delheure 1987: 395). Nous avons ainsi ici une opposition spatiale fondamentale *azayar* / oasis propre à ces communautés qui connaissent un mode de vie entièrement différent de celui des populations de montagne de l'Atlas marocain. L'oasis c'est le centre de vie et d'habitat qui s'oppose à *azayar*, espace

aride et sec. (Nous retrouvons ici la même conception géographique extérieur → désert que nous avons dans d'autres sociétés berbères. Il s'agit souvent des sociétés nomades sahariennes ou des populations des oasis. C'est l'exemple du terme *esuf* en touareg du Niger qui signifie "extérieur" et par extension "steppe déserte, solitude". En touareg de l'Ahaggar, le terme *ténéré* désigne "extérieur" par rapport l'habitat humain, et par extension "désert plat" (Foucauld 1951: 1397–1399). Le sens de l'extérieur est toujours perçu ici par opposition à l'habitat humain. (V. P. Galand-Pernet 1992; D. Casajus 1981).)

akerkur, takerkurt

Appartient à ce champ morphosémantique de « dureté, sécheresse » le terme *akerkur*, fém. *takerkurt* (réalisé parfois avec une pharyngalisation de r: *akerkur*, *takerkurt*). Ce terme est connu dans de nombreux parlers berbères avec le sens de « tas de pierres, endroit sec et pierreux » (tachelhit, Maroc central). Comme les consonnes vélaires *k*, *g*, *ɣ* et *q* connaissant de nombreuses alternances dans la phonologie diachronique berbère, on peut rapprocher *akerkur* d'autres termes comme *aqerqar*, *ayeryur*, *tayaryart*, ... attestés dans certains parlers: *aqerqar* « endroit sec et pierreux (kabyle, Dallet 1982: 670), *tayaryart*, pl. *tiyaryarin*, « endroit de terre séché et durci par le soleil » (touareg de Niger, Prasse et al. 1998: 121), *tiyeryert* « sol de maison » (kabyle, Dallet 1982: 623), *ayeryur* « élargissement de la vallée, à sol uni et dur, en plaine » (touareg de l'Ahaggar, Foucauld 1951: 1755).

Rappelons que de nombreux toponymes avec *akerkur* ou la forme du féminin *takerkurt* sont attestés dans plusieurs localités du sud-ouest du Maroc et dans les zones montagneuses du Haut-Atlas occidental (Laoust 1940). Ces toponymes désignent souvent des lieux et des terrains secs, arides et pierreux. Ajoutons que le massif montagneux central du pays des Kel-Ahaggar est connu chez les touaregs par le nom *Takerkurt n-Ahaggar* (Foucauld 1951: 875).

Par ailleurs, nous avons avec les exemples *akerkur*, *takerkurt*, *amquqr*, *aqerqar*, *ayeryur*, ... un procédé morphologique de redoublement consonantique bien connu en lexique berbère (Chaker 1990; El Mountassir 1996, 2002) ainsi que dans d'autres langues chamito-sémitiques. Ce mode de formation lexicale est d'origine expressive avec un sens fréquentatif et augmentatif. (P. Guiraud (1967) désigne ce type de formation lexicale, qui est connue aussi dans d'autres langues, par formation tautologique.) Les mots formés à partir de ce modèle morphologique traduisent souvent des réalités qui se caractérisent par la multitude, l'affluence et l'abondance. Ainsi, le terme *akerkur* (redoublement des consonnes *k* et *r*) « terrain pierreux / tas de pierres » doit être interprété comme « terrain avec des pierres ici et là; terrain parsemé de pierres ».

Ces mots appartiennent ainsi à un modèle morphologique qui est associé à un sens. En d'autres termes, nous avons ici un paradigme constitué d'un ensemble de mots qui ont des structures morphologiques analogues: le redoublement consonantique. Il s'agit dans la plupart des cas de redoublement des deux consonnes. Ce paradigme est un modèle à partir duquel de nouveaux mots peuvent être créés.

C'est dans ce type de formation lexicale que la notion de structure est fondamentale: une fois le mot trouve sa cohérence morphologique à l'intérieur de ce paradigme, il trouve en même temps sa cohérence sémantique. Rappelons qu'au niveau méthodologique, on ne part pas ici des mots isolés, mais d'un paradigme où les mots viennent trouver leur place. Le chercheur qui analyse le terme *takeṛkurt*, par exemple, se réfère à une série de mots comme *amquqr*, *tiṣeryert*, *ayeryur*, *yeryer*, etc., c'est-à-dire au paradigme d'où il tire sa signification. Une bonne partie du vocabulaire géographique berbère se caractérise par ce mode de formation lexicale. (C'est le cas par exemple de *tagragra*. Ce terme, qui appartient au vocabulaire géographique, désigne tout terrain ou territoire qui présente la forme d'une carapace *tagra*. Cela correspond aux endroits ballonnés (cuvette, bassin, etc.). Par ailleurs, il importe de signaler que cette forme *tagragra* est largement attestée en toponymie dans plusieurs localités du Souss.)

Pour ce procédé de formation lexicale, aussi fréquent et connu en berbère, il nous manque une étude avec un inventaire exhaustif qui permettrait d'expliquer un grand nombre de termes appartenant au vocabulaire géographique dont le sémantisme reste jusque-là obscur. Cette étude nous permettrait également d'avoir des listes de mots classés d'après leur mode de formation. (L'expression linguistique de la répétition d'un mouvement ou la multitude d'une réalité se fait dans d'autres langues par des unités lexicales à redoublement; voir à ce propos les travaux de F. Skoda (1982).)

Il est important de signaler aussi que dans certains parlers berbères le sens de « pierre, galet, caillou » tire sa signification à partir de la notion de « dureté, sécheresse » avec la racine KR. C'est ce que nous retrouvons dans le parler des Ayt Hdidou (Haut Atlas oriental): *akuray*, fém. *takurayt* « petit galet, petite pierre » : *iemmer wasif igran s-ikurayn* « la rivière a recouvert les champs de galets » (Azdoud 1985: 124).

Sécheresse, plage et aire à battre les céréales

Le mot *tayart*, attesté dans plusieurs parlers berbères, est le nom verbal de *yar* « être sec, dur », d'où le sens de « sécheresse, dureté, durcissement, aridité »:

tayart « sécheresse », tachelhit

tayart « terre sèche », tachelhit (Asaka, plaine de Tiznit)

tayart n'wanu « partie sèche du puits, fond sec du puits », tachelhit

tayert, *ayurar* « sécheresse, aridité », kabyle (Dallet 1982: 622)

tayart « action de sécher, d'être sec; séchage, durcissement, sécheresse, dureté », Ayt Hdidou (Azdoud 1985: 181)

tayart, *tayara* « sécheresse, aridité, dureté », Maroc Central (Taïfi 1991: 198)

tyuri, *ayari* « sécheresse », Beni-Snous (Destaing 1914: 321)

tayart, pl. *tayaren* « sécheresse », touareg de Niger (Prasse et al. 1998: 118)

teyart « fait d'être sec, de se sécher », touareg de l'Ahaggar (Foucauld 1951: 1752)

Ce même terme *tayart* prend un autre sens dans le vocabulaire maritime tachelhit. En effet, ce terme désigne dans cette zone géographique « plage », ce qui signifie

littéralement « bord de mer avec du sable sec ». Dans ce contexte, la série *yar* « être sec », *syer* « sécher », *asyar* « le fait de sécher », *tayart* « plage » forment un champ lexical appartenant au vocabulaire de l'espace maritime et de l'espace aquatique en général: rivière, cours d'eau, ruisseau, etc. J'ai relevé dans plusieurs localités de pêche au sud d'Agadir des termes et des expressions en rapport avec cette notion de « dureté, sécheresse » qui décrivent l'espace maritime (El Mountassir 2015). Nous avons ainsi le terme *asyar* -nom verbal de *syer* « sécher »- qui désigne dans cette zone littorale « marée basse ». Le moment où la mer commence à se retirer, cette situation se traduit dans la langue par l'expression *ar isyar lbher*, ce qui signifie littéralement « la mer fait sécher / se sèche ». En d'autres termes, l'espace du littoral est perçu comme un espace sec, découvert par la mer au moment de la marée basse (moment de reflux) d'où le terme *tayart* qui désigne dans ce contexte « plage ». Rappelons que ce même terme désigne en langage terrien « sécheresse ».

Dans certains parlers berbères, le nom de l'«aire à battre les céréales» est dénommé à partir de la notion de « dureté, sécheresse ». C'est l'exemple touareg *tiyeryert* « aire à battre le grain » = « terrain plat et dur pour battre le grain ». Ce terme fait partie de la même famille de mots que *yeryer* « rendre sec et dur » et *ayeryer* « fait de rendre sec et dur » (Foucauld 1951: 1751, 1753–1754). *tiyeryert* signifie en kabyle « sol de maison » (Dallet 1982: 623), et dans certains parlers ce même terme prend le sens de « surface plane et bien damée ».

Lexique géographique berbère: une diversité de désignations et une unité de signification

L'ensemble de ces termes d'origines si diverses et si dispersés dans l'espace forment bien une homogénéité frappante. L'idée initiale ou le protosémantisme qui supporte l'ensemble est celle de « dureté, sécheresse ». Tous les termes de ce champ mettent en évidence les traits pertinents décrivant une réalité spatiale qui se caractérise par la dureté et la sécheresse.

Il s'agit d'un bel exemple qui illustre ce contraste entre l'unité du vocabulaire géographique berbère et les multiples variations dialectales. Si les divers parlers berbères ont développé des systèmes de désignations assez variés dans le domaine du vocabulaire géographique, le fonds commun de ce vocabulaire reste reconnaissable. Cette diversité lexicale reflète celle des modes de vie et de l'organisation sociale du monde berbère. Le vocabulaire relatif à l'espace géographique s'adapte avec le milieu naturel et social et varie avec lui. Ceci est évident: la terminologie de l'espace n'est pas la même chez les touaregs du désert que chez les communautés sédentaires du sud-ouest du Maroc.

Dans ce sens, il est fort intéressant de voir comment ce changement social et institutionnel s'est opéré dans la langue berbère selon les régions. Car, il est important de rappeler ici que les diverses régions du monde berbère n'ont pas connu les mêmes évolutions sociales. Les sociétés berbères sont restées isolées les unes des autres si bien que chaque région possède ses spécificités linguistiques. Il est importe de préciser ici que la langue est un fait social, et comme tous les autres faits

sociaux, elle évolue avec le temps et se diversifie dans l'espace. C'est dans ce contexte que certains linguistes parlent d'étude géologique des mots qui consiste à situer chronologiquement et géographiquement leurs significations, à définir leurs rapports et à reconstituer leur genèse.

A travers cette analyse de la base YR , j'ai tenté de saisir quelques mécanismes d'évolution et d'adaptation du vocabulaire spatial de « dureté, sécheresse » tout en tenant compte de la dimension géographique et sociale que véhiculent les unités lexicales. Il importe de rappeler dans ce sens, que cette analyse s'inscrit dans une approche qui étudie la langue avec tout l'univers humain qui l'entoure. Nous ne pouvons pas par exemple saisir la parenté entre les termes *yar*, *azayar*, *tayart*, *takerkurt*, *aqerqar*, *ayeryur* etc., sans tenir compte du milieu géographique et social dans lequel ces mots sont attestés. En d'autres termes, cette approche donne la priorité à la dimension anthropologique de la langue, sachant que celle-ci n'est jamais une entité disjointe de la réalité et du vécu humain. D'autre part, cette analyse montre que, dans les travaux sur le lexique, il faut renoncer aux méthodes qui ne considéreraient les mots qu'en termes isolés, ne tenant pas compte de leur champ dérivationnel avec la totalité des acceptions sémantiques. Traiter les mots dans leur série dérivationnelle la plus large possible permet de révéler les spécificités d'une langue et sa vision du monde.

C'est dans ce sens que nous voulons montrer dans cette analyse que la dérivation n'est pas conçue ici comme une simple juxtaposition des termes, c'est-à-dire de simples évolutions mécaniques; mais que les dérivés illustrent des opérations mentales et témoignent d'une vision du monde propre à une communauté donnée. A ce propos, A. J. Greimas (1966: 75) estimait que la dérivation « constitue dans une large mesure, un système de classification ». Ainsi, c'est en étudiant les séries dérivationnelles des mots d'espace géographique en rapport avec leur noyau de base de « dureté, sécheresse » que nous pouvons appréhender la manière dont les locuteurs berbères perçoivent ce type d'espace.

La séquence bi-consonantique YR , support de l'idée initiale de « sécheresse, dureté », constitue le signifié de puissance qui a engendré plusieurs séries de mots avec le sens de « lieu, endroit sec ou dur (terrain pierreux et inculte; espace désertique, ...) ». Le passage de signifié de puissance aux mots s'effectue par un certain nombre de dérivations morphologiques et sémantiques comme nous l'avons vu dans les exemples précédents. Entre le mot *tayart* « plage » par exemple et le signifié de puissance YR « dureté, sécheresse », la dérivation est à la fois morphologique et sémantique: morphologique, car tous deux possèdent un signifiant commun *yr-*; sémantique dans la mesure où le mot *tayart* est engendré à partir de sèmes « espace + du sable sec au bord de mer ».

Par ailleurs, l'étude de ce champ morphosémantique montre aussi qu'un même mot peut avoir plusieurs significations selon les parlers. La démarche consiste à établir le noyau sémique commun à ces différentes significations. Il s'agit ici d'une opération diachronique (et historique), car ces termes appartiennent à des aires géographiques diverses et sont attestés à des périodes linguistiques différentes. Le cas de *azayar* « endroit sec, plaine sèche, terrain boisé, pâturage desséché, extérieur, désert, ... », est un bel exemple qui illustre ce phénomène. C'est dans ce contexte

que nous adoptons la distinction fondamentale formulée par Benveniste (1969: 12) qui consiste à ne pas confondre signification et désignation. Pour ce linguiste, il faut savoir «établir certaines distinctions essentielles [...] entre désignation et signification, à défaut de laquelle tant de discussions sur le sens sombrent dans la confusion. Il s'agit, par la comparaison, et au moyen d'une analyse diachronique, de faire apparaître une signification là où, au départ, nous n'avons qu'une désignation».

Pour ce même linguiste, une bonne méthode linguistique comparative dans le domaine du vocabulaire consiste à chercher, au-delà des désignations, à atteindre le niveau profond des significations. Dans ce sens, chaque champ morphosémantique présente une remarquable diversité de désignations, mais qui correspondent, à un niveau plus profond, à une unité de signification.

Références bibliographiques

- Azdoud D (1985) Lexique et textes des Aït Hadiddou (Maroc Central). Thèse de 3ème cycle, Sorbonne EPHE et Paris III
- Benveniste E (1955) Homophonies radicales en indo-européen. *Bulletin de la Société Linguistique de Paris* 51:14–41
- Benveniste E (1969) Le vocabulaire des institutions indo-européennes.2. pouvoir, droit, religion. Les éditions de Minuit, Paris
- Bernus E (1982) Points cardinaux: les critères de désignations chez les nomades touaregs et maures. *Bulletin des études africaines de l'INALCO* I(2):101–105
- Casajus D (1981) La tente et le campement chez les touaregs Kel Ferwan. *Revue des mondes musulmans et de la Méditerranée* 32(2):53–70
- Chaker S (1979) Les racines berbères trilitères à 3^{ème} radicale alternante. *Comptes-rendus du G.L.E. C.S XVIII–XXIII(fasc.2i)*:293–303
- Chaker S (1990) Comparatisme et reconstruction dans le domaine chamito-sémitique: problèmes de méthodes et de limites. *Travaux du Cercle Linguistique d'Aix-en-Provence* 8:161–186
- Chaker S (1995) Linguistique berbère. Etude de syntaxe et de diachronie. Peeters, Louven
- Cohen D (1975) Rapports sur les conférences de Sémitique comparé. *Annuaire 1974/1975 de l'Ecole Pratique des Hautes Etudes* 265–275
- Dallet JM (1982) Dictionnaire kabyle-français. Parler des Aït Mangellat. SELAF, Paris
- de Foucauld Ch (1951) Dictionnaire touareg-français, dialecte de l'Ahaggar, 4 tomes. Imprimerie Nationale, Paris
- Delheure J (1984) Dictionnaire mozabite-français. SELAF, Paris
- Delheure J (1987) Dictionnaire Ouargli-français. SELAF, Paris
- Destaing E (1914) Dictionnaire français – berbère (Dialecte des Beni- Snous). Ernest Leroux, Paris
- El Mountassir A (1996) Etude d'un champ morphosémantique en berbère: le cas de l'étymon BR. *Awal* 14:81–87
- El Mountassir A (1998) La perception de l'espace en tachelhit. In: Cas de présaharien marocain, Actes du colloque Les oasis du Bani, Tata 15–17 mai 1995. Publications de la Faculté des Lettres et des Sciences Humaines d'Agadir, Agadir, pp 125–138
- El Mountassir A (2001) Espace, langue et société chez les Ichelhiyn. Sud-ouest du Maroc. Thèse de Doctorat d'Etat, Université Ibn Zohr, Faculté des Lettres et des Sciences Humaines, Agadir
- El Mountassir A (2002) La notion de “mou, humide, tendre” en berbère: l'étude de la base LY. In: Naït-Zerrad K (ed) *Articles de linguistique berbère, Méorial Werner Vycich*. L'Harmattan, pp 181–196
- El Mountassir A (2005a) Lexique et perception de la réalité en berbère. L'exemple du terme *azayar*. *Faits de langues* 1(26):259–269

- El Mountassir A (2005b) Kerkur. In: Chaker S (ed) *Encyclopédie Berbère*, vol XXVII. Peeters, Louven, pp 4176–4178
- El Mountassir A (2010) Maison. In: Chaker S (ed) *Encyclopédie Berbère*, vol XXX. Peeters, Louven, pp 4525–4528
- El Mountassir A (2015) Ces voix et ces mots qui disparaissent. Le vocabulaire maritime berbère (tachelhit). *Revue des Etudes Berbères* 10:189–203
- Galand L (1997) A propos des noms de l'écorce. *Rocznik Orientalistyczny (Orientalia Memoriae Professoris Thaddaeo Lewicki Dedicata)* 50(2):95–102
- Galand-Pernet P (1963–1966) Ver, coucou, coup de takouk. Notes sur un champ morphosémantique en berbère. *Compte rendus du GLECS* X:6–14
- Galand-Pernet P (1984) Sur quelques bases radicales et champs morphosémantique en berbère. In: Bynon J (ed) *Current Progress in Afro-Asiatic Linguistics (Papers of the Third International Hamito-Semitic Congress)*. John Benjamins, Amsterdam, pp 291–303
- Galand-Pernet P (1985) Blanc, lumière, mouvement. *Littérature Orale Arabo-Berbère* 16/17:3–20
- Galand-Pernet P (1992) Du désert-vie au clos construit: espaces berbères. *Revue d'études islamiques* LV/LVII:381–406
- Greimas AJ (1966) *Sémantique structurale*. Larousse, Paris
- Guiraud P (1967) *Structures étymologiques du lexique français*. Larousse, Paris
- Haudry J (1985) La reconstruction. *La linguistique* 21:91–107
- Laoust E (1920) Mots et choses berbères. *Dialectes du Maroc*. Challamel, Paris
- Laoust E (1932) Siwa: son parler. Leroux, Paris
- Laoust E (1940) Contribution à une étude de la toponymie du Haut- Atlas. *Revue d'études islamiques* 1/2:27–77
- Meillet A (1925) *La méthode comparative en linguistique historique*. Champion, Paris
- Mounin G (1975) *La sémantique*. Seghers, Paris
- Naït-Zerrad K (2000) Autour de la base morphosémantique SK en berbère. In: Chaker S, Zaborski A (eds) *Etudes berbères et Chamito-sémitiques, Mélanges offerts à Karl-G. Prasse*. Peeters, Louvena/Paris, pp 295–302
- Picoche J (1993) Analyse lexicale et perception de la réalité. *Cahiers de praxématique* 21:37–52
- Prasse KG, Ghoubaid A, Ghabouane M (1998) *Lexique touareg – français*. Museum Tusculanum Press, Université de Copenhague [Carsten Niebuhr Institute Publications], Copenhagen, p 24
- Skoda F (1982) Le redoublement expressif: un Universal linguistique. *Analyse du procédé en grec ancien et en d'autres langues*. SELAF, Paris
- Taïfi M (1991) *Dictionnaire tamazight-français (parlers du Maroc central)*. l'Harmattan-Awal, Paris

Le vocabulaire berbère des plantes. Profondeur historique, conservation et permanence

33

Mohand Tilmatine

Contents

Introduction	699
Les sources	700
Les difficultés spécifiques aux sources anciennes et modernes	701
Les sources contemporaines	702
La sémantique du vocabulaire des plantes	703
Aspects phonétiques	705
Structures composées	706
<i>Composés avec bu + X</i>	706
<i>Composés avec substantif+substantif</i>	707
<i>Verbe + Substantif</i>	707
<i>Composés avec tif/tuf/f/af</i>	707
<i>Composés avec eilal “herbe”</i>	708
<i>Composés avec mer-</i>	709
<i>Composés avec wa-</i>	710
<i>Composés avec -enni,-enna, -emmi</i>	710
Archéologie linguistique et récupération-reconstruction lexicale	711
Les emprunts	712
Conclusion	714
Bibliographie consultée	716

Introduction

Cette contribution abordera le vocabulaire des plantes du point de vue de sa profondeur historique, de son maintien ou disparition, mais aussi pour en dégager certaines caractéristiques sémantiques et surtout morphologiques avec un intérêt particulier pour le phénomène de la composition.

M. Tilmatine (✉)
University of Cádiz, Cádiz, Spain
e-mail: mohand.tilmatine@uca.es

Cette contribution se basera fondamentalement sur un lexique berbère dégagé dans le cadre d'un travail d'édition mené sur un manuscrit de botanique du XI–XII^e siècle (Réédition de la *'Umdat at-ṭabīb fī ma'rifat an-Nabāt* attribuée au Sévillan (Abū l-Ḥayr al Iṣbīlī), par Federico Corriente (université de Saragosse), Joaquín Bustamante et Mohand Tilmatine (Université de Cádiz). Edition du texte arabe (2004), traduction commentée à l'espagnol (2007) et des annexes et différents indexes onomastiques en deux volumes (2010). Cette édition sera citée comme BCT- I, II ou III vol. 1 et III vol. 2, suivie du numéro de la page en question.), réalisé entre 1998 et 2010 et qui s'est clôturé avec une nouvelle édition du traité de botanique d'al Iṣbīlī - qui comprendra l'étude détaillée de deux cent soixante-douze (272) phytonymes amazighs (BCT- III: vol. 2, 926–928). Une première liste des phytonymes berbères avait été établie en 2002 (Tilmatine and Bustamante [=T&B]).

L'exploitation de ce manuscrit permet de reprendre certaines réflexions sur la phytonymie berbère, exprimées dans certains travaux, parmi lesquels l'ouvrage de Laoust (1920), tient, malgré son ancienneté, une place prépondérante et qui demeure un des plus importants et des plus fournis dans les études berbères. (Nous utiliserons dans ce texte souvent le terme « berbère » qui apparaît dans l'original en arabe tout en étant conscients de la valeur symbolique de l'usage de l'autoglottonyme équivalent « amazigh ».)

L'étude sera précédée par la présentation des principales sources anciennes et contemporaines pour l'étude des phytonymes amazighs.

Enfin, ce travail ne saurait prétendre à aucun type d'exhaustivité, ni en ce qui concerne le vocabulaire recensé, ni en ce qui concerne les catégories d'analyse utilisées. L'objectif est davantage de proposer une approximation à la problématique des phytonymes amazighs dans les textes anciens et de leur rôle comme fonction d'appoint aux efforts actuels en vue de développer la langue amazighe dans une phase aussi cruciale que la présente avec son passage d'un état de langue oral à celui de langue écrite. La transcription des exemples suit celle des différents auteurs cités.

Les sources

Le traité de botanique andalou, *'Umdat at-ṭabīb fī ma'rifat an-nabāt*, est un ouvrage unique à bien des égards. Il s'agit d'un recueil détaillé de toute la botanique nécessaire en son temps pour l'exercice de la médecine, présenté selon un système de classification original, bien développé. Il regroupe les plantes en genres et espèces et fournit des descriptions botaniques qui dénotent chez son auteur d'excellentes capacités d'observation. L'ouvrage constitue une source de première importance pour l'étude de la botanique médiévale en al-Andalus. En outre, un volume considérable de lexique phytonomique est lié, non seulement à l'arabe, mais aussi à l'amazigh ainsi qu'à d'autres langues comme le grec, le roman, le persan, le syriaque etc. Le lexique roman a été étudié par Miguel Asín Palacios (1943), qui n'a édité que quelques passages basés uniquement sur le manuscrit de l'Académie royale d'histoire de Madrid. L'édition complète, selon les deux manuscrits qui nous sont parvenus, le numéro 3505 ٣ de la Bibliothèque générale de Rabat (*al-Ḥizāna*

al-’amma li-l-Kutub wa-l-Watā’iq), et le numéro 40 (XL) de la bibliothèque de l’Académie royale d’histoire de Madrid, n’est apparu qu’en 1990 de la main de al-Ḥaṭṭābī (1990). Ce dernier attribue la paternité de la *’Umdat aṭ-ṭabīb* à l’agronome sévillan Abū l-Ḥayr al-Isbīlī. La question de savoir s’il s’agit réellement de son auteur ou non est néanmoins parfois remise en cause (Carabaza Bravo 1995). Benllakhdar (2010: 12), quant à lui, émet l’hypothèse selon laquelle al-Isbīlī aurait reçu ses informations d’un “proche du Roi al-Mu’tamid b. ‘Abbād qui l’aurait accompagné durant son exil, ou bien, par “un voyageur averti, qui avait l’habitude de faire des allers-retours dans ce pays”, mais l’examen de l’œuvre elle-même permet d’extraire des données qui confirment son autorité en matière de lexique amazigh et qu’il en est bien l’auteur. D’après ce qu’il dit lui-même, il semble clair que l’auteur aurait vécu à Séville durant la dernière période du règne d’al-Mu’tamid ibn ‘Abbād et qu’il connaissait bien Ibn Baṣṣāl et Ibn al-Lunquh – ce dernier est même cité comme son “grand maître” (BCT- II: 513). Les deux fameux botanistes de Tolède, qui, après la chute de cette ville aux mains des Castillans en 1085, ont été liés au Jardin royal de Séville. La date de la composition de l’œuvre doit être au moins la première moitié du XIIe siècle, car elle parle d’Ibn al-Lunquh comme étant déjà décédé, et on sait qu’il est mort en 1104 (Bustamante and Tilmatine 1999: 46).

Les difficultés spécifiques aux sources anciennes et modernes

Au-delà de ces problèmes généraux liés directement aux caractéristiques linguistiques de l’amazigh, des difficultés extralinguistiques rendent encore plus ardu le travail de compilation, d’agencement et d’organisation du lexique des phytonymes amazighs.

Les sources anciennes concernant le lexique des plantes renvoient surtout aux manuscrits de botanique arabes. La transcription de l’amazigh dans ces manuscrits suivait un certain nombre de règles, que nous commençons à connaître de mieux en mieux (Van den Boogert 1989). Cette transcription n’est cependant pas la même dans tous les manuscrits, d’où les difficultés de déchiffrement.

D’autres erreurs peuvent être également motivées par des critères purement graphiques. Ainsi, en alphabet maghrébin, le *qāf* s’écrit avec un point souscrit. L’on peut imaginer les risques de confusion, surtout en position médiane où les lettres *qāf* (ق), *ghayn* (غ), *fā’* (ف) et *’ayn* (ع) se ressemblent beaucoup et peuvent facilement induire à l’erreur.

Contrairement aux manuscrits plus récents (à partir du XVI–XVII^e s.) qui sont en général complètement vocalisés, le texte de la *’Umda*, ne l’est pas toujours et rend plus difficile la lecture des termes inconnus. L’intervention des copistes qui peuvent falsifier des éléments lorsque ceux-ci estiment qu’une forme est plus correcte, est un autre facteur dont il faudrait tenir compte.

Signalons enfin que la transcription est très souvent approximative, souvent hasardeuse, car les auteurs ne sont pas, en général, berbérophones. Ils travaillaient donc avec des informateurs et s’efforçaient de reproduire ce qu’ils pensaient avoir entendu.

Du point de vue de la collecte des phytonymes, il faut ajouter qu'ils sont le plus souvent éparpillés dans le corpus du texte; ce qui rend plus laborieux le travail de recherche et de regroupement de ce lexique.

Par ailleurs, certains termes, bien qu'amazighs, sont tellement bien implantés dans la langue qu'ils passent parfois pour arabes (cas de l'**addad** "chardon à glu" (*Atractylis gummifera*). L'origine amazighe de ce terme échappe même à de grands spécialistes comme Ahmed Issa, auteur pourtant d'un ouvrage de référence sur les phytonymes en plusieurs langues et qui se contente de citer **addād** comme un des noms de l'*A. gummifera* à côté d'autres termes comme *iṣḥīṣ* (arabisation du grec *Ixia*) ou *ṣūk al-'ilk*. (BCT- I (86) et BCT-II (138) donnent son origine berbère, en revanche des auteurs comme Ahmed Issa (1930: 27) donnent pour l'*A. gummifera*, **addād** et également le mot *al-wahīd* pour le Maghreb. Ni Issa, ni Trabut (1935: 42) mentionnent qu'il s'agit d'un terme berbère alors que des auteurs d'ouvrages sur le berbère l'identifient en tant que berbère comme tel (Chafik: TOME I, 585; Laoust 1920: 509; Boogert 1998: 178; Dozy TOME I: 14, etc.).)

Les sources contemporaines

Elles ont en commun le fait de pas constituer des ouvrages monographiques dédiés à la botanique amazighe, tout en intégrant un nombre plus ou moins important de termes amazighs dans leurs listes terminologiques. Les données amazighs sont considérées dans ce genre d'ouvrages comme accessoires.

Nous pourrions répartir ces ouvrages contemporains *grosso modo* en deux blocs:

- a-. Les ouvrages ou articles, très peu nombreux par ailleurs, qui, comme le travail de E. Laoust (1920, dorénavant LM), ont pour objet l'amazigh en général. Ces ouvrages, souvent de type ethnographique, contiennent un certain nombre d'informations, parfois très importantes sur le lexique des plantes. L'ouvrage de Hanoteau-Letourneux (1893, dorénavant [HL]) et d'autres titres comme ceux de Trabut (1935), Salmon (1906), Renaud and Colin (1934) ou Mercier (1905) pourraient également entrer dans cette catégorie.
- b-. Ensuite des ouvrages de botanique qui portent sur la phytonymie en Afrique du Nord et qui donc, du fait de la forte présence de l'élément amazigh contiennent inévitablement la terminologie amazighe correspondante.

Depuis quelques années, nous assistons à la publication de quelques ouvrages de phytonymie de la main de Nord-africains.

Les auteurs de ces derniers ouvrages sont en général des pharmacologues ou des botanistes. (Voir en particulier le travail de Jamal Bellakhdar 1997 [Dorénavant BELL].) Ils veulent donc faire avant tout des livres sur les propriétés des plantes, mais pas un catalogue ou des lexiques plurilingues. Ceci d'autant que souvent, ils ne bénéficient que de peu de connaissances linguistiques et encore moins de l'amazigh. Le résultat s'en fera sentir au niveau de la transcription, mais également au niveau des listes des mots avec une grande déformation des noms en amazigh ou

en arabe dialectal. Ces erreurs seront d'autant plus difficiles à corriger que ces auteurs ne disposent pas non plus de références lexicographiques avec, p. ex. des formes standardisées des différents phytonymes. Bien entendu, il y aura également très peu de détails sur l'appartenance ou origine dialectale des phytonymes concernés.

La sémantique du vocabulaire des plantes

Les références au monde animal sont très nombreuses. En cela, le lexique amazigh des plantes est organisé selon des structures comparables à celles d'autres langues (Moïnfar 1988). Ces aspects ne sont que des exemples du travail qui pourrait et devrait être entrepris dans le champ du lexique amazigh des plantes. Un terrain qu'avait commencé à défricher E. Laoust (1920). Ainsi, il fournit des éléments intéressants pour la connaissance de ce vocabulaire en analysant des composés divers soit avec des morphèmes préfixés ou en faisant des propositions quant à l'étymologie de certains noms de plantes.

Un bref regard sur les structures sémantico-formelles de ce champ lexical suffit à l'observateur pour se rendre compte rapidement que le lexique amazigh des plantes est articulé autour d'un système de comparaison du monde des plantes (aspect général, forme de la plante, de la feuille ou de la tige...) avec l'être humain, les animaux (Cf. par exemple *Patte-de-chat* et *œil-de-rossignol* de Mohammad Djafar Moïnfar (1988).), la religion, la mythologie et les croyances populaires. Voyons quelques exemples:

Être humain

Des termes tels que "vieille, vieillard, barbe, cheveux, dent, sein, oreille, berger" etc. peuvent servir de support de comparaison:

- **tughmas n temgharin** "dent des vieilles", **uriḍ umeksa/turiṭ** "Pet de berger" (HL: 104) "pissenlit" *Hyoseris radiata* (HL: 104), *Taraxacum dens-leonis*; *Osyris alba* L. **warneger** "celui qui laisse des enfants males, des rejetons" (HL: 117); **tammart n wemghar**, litt. "Barbe de vieillard", mousse de chêne, mousse de cèdre, *Evernia prunastri* Ach. (BELL: 6); **ifadden n tmurt**, orobanche, litt. "Jambes du pays" (Dallet 1982 [DAL]: 191), etc.

Références Religieuses

Comme il est facile de l'imaginer, le vocabulaire berbère des plantes a recourt aussi aux symboles religieux et aux croyances qui marquent la culture amazighe: **taṣeṭṭa n Meryem**, l'absinthe (Dallet 1963 = [FDB]: 19) [litt. "Branchettes de Marie"] qui s'appelle d'ailleurs également ainsi en arabe dialectal: **šajret Meryem** (BELL: 68); **taxlult n nbi (tixlulin n nnbi)** Narcisse (DAL: 896; Trabut 1935: 170); **xizzu n widayen**, la moutarde blanche, Aït Atta, [litt. "carotte des Juifs"] (LM: 512); **tarumit** (LM: 422), figue de Barbarie [litt.: "la Chrétien ne"] (La traduction « chrétien » pour « **abelluḍ** » n'est peut-être pas toujours correcte. En effet, même

si, cette traduction est référenciée comme telle dans les sources citées (Laoust 1920: 422 et 499), il n'en demeure pas moins que ce terme est souvent utilisé par opposition aux variétés locales. Ainsi, pour **urumi** (FDB: 11), le même Dallet dans son dictionnaire (DAL: 91) le traduit comme “importé”; “Fleur d’orme, d’acacia. • *yella ççilmum n tulmatin, n tmurt, d awray, yella ççilmum urumi d acebhan*, il y a la fleur d’orme, du pays, qui est jaune, et celle de l’acacia importé, qui est blanche”. Dans un autre exemple, sous la racine **urumi**, (DAL: 725), l’auteur nous donne pour **irumyen** (i) “Européen, spécialement Français. • *ad iru wi llan d arumi!* même un Européen pleurerait! • *abellud urumi* châtaignes (glands d’Europe)”. Même cas dans l’exemple “**ttejra ieedmen afriwen at tyid wi llan d arumi** un arbre sans feuille cela fait pitié aux moins sensibles (m. à m.: à un Européen)” (DAL: 604). Relevons que pour le sens “sylvestre”, on utilise souvent **n lexla** (des champs), **n wuccen** (du chacal), (*n*) *ughyu!* (de l’âne.); **ttejra urumi**, acacia [litt. “arbre du Chrétien”] (FDB: 11); **tigzinin urumi**, ortie romaine, litt. “du Chrétien”, “Kugelbrennessel”, *Urtica pilulifera* (TOPPER: 232), etc.

Mythologie

De figures de la mythologie amazighe comme l’ogre, apparaissent dans les noms de plantes sous différentes variantes: **agursal n tseryel**, litt. “champignon de l’ogresse” (LM: 502); **imezzughen n tseryel**, litt. “oreilles de fée” (FDB: 11); **iqšuşen n tgrođ** “parure de l’ogresse”, **Imeghran**: érodium (LM: 502); **taryal** “ogresse”, *Mandrogora autumnalis*, Mandragore (T&B: 443), etc. . .

Noms d’animaux

différents animaux forment la base des désignations de grand nombre de plantes: **ađar utbir**, “patte de pigeon”, fausse bourrache, A. Baâmran (LM: 500); **ađar n ts курт**, “patte de perdrix”; **aghembub n usu**, “bec de cigogne”, érodium, Ntifa (LM: 501); **idaren n tmušša [idaren]** “pattes des chats”, érodium, I. Oukensous (LM: 501); **ajdiđ n ileghman** “gale des chameaux”, ortie, Tlit; **duj ilughman**, “noix des chameaux”, chardon à feuilles panachées, **asennan ir’aman**, “épine des chameaux”, ronce, Ait Baâmran; **ibaun n iraman [ibawen ileghman]**, “fèves des chameaux”, légumineuse, Ihahan (LM: 501); **tikdert b wegħyul**, “l’épi des ânes”, *Hordeum Caput medusae* Coss et D. R. (HL: 129); **timžin b wegħyul**, “orge des ânes” *Hordeum murinum*, “orge des rats” (Trabut 1935: 130); **xizz[u?] izgaren**, “carotte des boeufs”, oseille?, Ida u Tanan; **axo n izgaren**, “lait des boeufs”, euphorbe, Warzazat; **tighmas ugennuz [ugenduz]** “dents de veau”, ind., Ichq.; **iles ugenduz**, “langue de veau”, vipérine, Zwawa; **isk uzgır**, “corne de boeuf”, id, Illalen; **tadut bulli [tađut n wulli]** “laine des brebis”, doucette, Kabylie; **aselğag n izamaren** “gomme des agneaux/moutons”, *Otanthus maritimus* L. Hoffmanns and Link (BCT-II, 604: Chafik TOME I: 291; et TOME II: 144; **aghrum n tgheţten**, “pain des chèvres”, champignon, A. Messad (LM: 501); **tarubia n tgheţten**, “[rubia] des chèvres”, garance, *Rubia peregrina* (LM: 520); **himessi wairad, [timessi wayrad]**, “feu du lion” Chenoua (LM: 501); **ađil ilugmaden**, “raisin des serpents”, plante grimpante, Ait Issafen (LM: 501); **Tilt tfighra**, *Trapa natans* L. “herbe du serpent” “chataigne d’eau” (T&B: 449); **aselen g iddaun [aslen n yiddawen]**, “frêne des singes”, *Daphne laureola*, Zwawa (LM: 501);

Tilt uriği [tilt urilli] “Phalangère ramifiée” *Anthericum ramosum*; *Anthericum liliago* (T&B: 449; Chafik TOME I: 408).

Souvent ces comparaisons portent sur certaines parties ou des membres d’animaux: **tādarin iigḍaḍ [tiḍarin igḍaḍ]**, litt. “pattes d’oiseaux”, Ida Gounidif; **tiḍarin iskuran [tiḍarin]**, litt. “Pattes de perdrix”, Aït Baâmran (LM: 503); **timghilt izimer**, “queue de béliet”, réséda, Achtouken (LM: 501); **aberdud izem**, “queue de lion”, panais, Ichqern, Berabers du Moyen Atlas, (LM: 500); **iles ugenduz**, “langue de veau”, vipérine, *Echium plantagineum*, kabyle (HL: 109); **tiṭ n uglu**, “oeil de grenouille”, euphorbe, Tazarin (LM: 501); **tameẓẓugh n tili**, “oreille de brebis”, papillonacée, igliwa; **ameẓẓugh n tili**, “oreille de brebis”, renoncule, Tlit; **ameẓẓugh n uskai [uskay]**, “oreille de lévrier”, réséda, Ait Baâmran (LM: 501).

La langue recourt également à certains animaux sauvages ou domestiques pour désigner les plantes sylvestres: **ibawen ggilef**, “Fèves du sanglier”, lupin, *Phaca baetica* L. (HL: 85); “Oreille du sanglier”; **aḍil bbuššen**, “les raisins de chacal”, belladone (un sédum), Mtougga; **tiḍarin wwuššen** “pattes de chacal”, érodium, Ida Oukensus (LM: 502); **tarubya n (t) ghetten** garance *Rubia peregrina* “rubia des chèvres”, Zemmour (LM: 520); **lbşel iḍan** “oignon des chiens”, id. scille, Zemmour; **lbruaj iḍan** “asphodèle des chiens, scille maritime”, Aït Seghrouchen (LM: 501); **azir n wegħyul** “romarin des ânes/sylvestre”, (T&B: 425).

Parfois, cependant on recourt pour les espèces sylvestres tout simplement à des expressions comme **Lexla** “champêtre”: **I-latay I-lexla** “thé des champs”; **ssa’a I-lexla** “graine d’érodium”; **titshuffidin I-lexla** “feuilles de l’umbilicus”.

D’autres exemples peuvent étayer les observations de Laoust. Ainsi les constructions avec base **tf** dans le sens “être mieux que, surpasser” se retrouvent dans plusieurs cas comme: **tif el ḥenna** (T&B: 447); **tifuzzal** “le cyste”(FDB); **tuf ṭulba** “ivette, meilleure que les médecins” (BELL: 277) et probablement **tif sighra**, pour le moment non identifié (“genêt du Sahara”? *Onobrychis viciaefolia* Scop., T&B: 266), qui peuvent s’ajouter à la liste de Laoust, qui cite entre autres **tufaman** “plante aquatique; littéralement “mieux que l’eau”, Aït Atta; **tuf agho** “meilleure que le lait” (LM: 494).

Aspects phonétiques

Au plan phonétique et phonologique, l’étude des noms de plantes – dont certains sont très anciens - reflète la situation actuelle sans profondes modifications. Nous retrouvons ainsi certaines réalisations spécifiques à quelques dialectes telles que nous les connaissons aujourd’hui. Nous en donnons quelques exemples:

- a) - Affaiblissement du t > h (Aurès, Chenoua...): **taraghla** > **haraghla** (Chenoua); **tafruit** “glaieul”, Ill., Aït Baâmran; **Hafruit uuššen**, [tafruit n wuccen], litt. “épée/couteau du chacal”, Chenoua; **himessi wairad [timesi n wayrad]** “feu du lion”, Chenoua; **taselgha** > **haselgha** (Trabut 1935: 121).
- b) - l < r; l > n; ll > dj; ll > nn: **ilel** “laurier-rose” AHAGGAR (Tuareg); **alili** Sous nous donne **ariri** dans le Rif; **alili** > **anîni**, Aït Atta (LM: 503), (Cette

assimilation $l > n$ est relevée également par Kossmann pour le parler de Figuig (1997: 22).) **alidji** Izayan; **aghrum n unni** “pain des brebis”, champignons, Aït Atta [pour **aghrum n ulli**].

c) - g > y **agerni** > **ayerni** (LM: 513) “arum”, *Arisarum vulgare*.

d) Relevons également des phénomènes d’apophonie, donc d’une modification vocalique dans un lexème en passant d’une langue berbère à l’autre, tout comme d’ailleurs, elle peut exister au sein de la même langue. Le changement peut concerner une ou plusieurs voyelles. Cela peut toucher une voyelle initiale, une voyelle finale ou une voyelle interne. Exemple d’apophonie a/i: **Admam** Aubépine *Crataegus laciniata* Ucr. (Moyen Atlas, Haut Atlas, Rif, BELL: 441), > **idmim** kabyle (DAL: 143); **ilili** “laurier-rose” (kabyle) > **alili** (Tach., DAL: 167).

Structures composées

Les noms de plantes montrent que contrairement à ce que l’on pourrait penser la composition joue un rôle non négligeable en berbère. Bien sûr, l’absence d’ouvrages de référence en matière d’étymologie, voire lexicographiques, la connaissance encore limitée des variantes berbères ainsi que l’absence de travaux comparatistes interdialectaux rendent difficile le travail de reconstruction ou d’identification des éléments constitutifs de certains termes.

Ainsi, le champ de la composition et de l’identification des phytonymes composés demeure largement ouvert.

Le travail de Laoust illustre bien la situation à cet égard. Abordant cet aspect (LM: 493), il suppose cette juxtaposition, ou il en a l’intuition, sans toutefois toujours expliquer ou commenter la composition des deux parties du syntagme.

Un certain nombre de composés sont des éléments connus et relativement bien identifiés:

Composés avec bu + X

Certains noms de plantes sont composés avec **bu+X** et leur correspondants au féminin **m-**, **mu-**, ou **ma-** dans le sens: “maître de . . ., possesseur de . . ., producteur de . . .”: **bu-anzaren**, “saugé à deux couleurs”, Zwawa, **bu + anzaren** “nez, narines”; **bu-isennan** (en), “oseille sauvage”, Ouameslakht, **bu + sennanen** “épinés, piques, ronces”; **bu-zegzaw** litt. “Celui du vert/bleu”, Chardon-marie *Silybum marianum* L. (BELL: 108); **mmoghi**, littéralement “celle du lait”, “euphorbe”; **muxu** [*mmuxu*], “laiteron”, (LM: 494).

Les cas de composition **bu + verbe** sont moins productifs et moins courants, mais peuvent apparaître comme dans **bu-qsas** “ortie”, composé de **bu** et du verbe **qqes** qui signifie dans la majorité des dialectes “piquer, faire mal” > “ce/celui qui fait mal”. Un phytonyme ancien, aujourd’hui, semble-t-il disparu, est cité clairement comme “berbère” dans la ‘Umda: *Buriqdas* “Statice commun” ou “Lavande de mer”,

Limonium vulgare Miller (T&B: 431). Le terme est analysé par les auteurs comme une composition de **bu+areqqaq+idis**, littéralement “celui aux côtés fins”.

Composés avec substantif+substantif

Laoust (LM: 492 et suiv.) donne une série de termes composés avec différents éléments: **axu/agh**, “lait” : **tanaghut** (Composé des éléments suivants: premier élément du circonfixe pour marquer le genre féminin **ta-** + **n** (marque du génétif (=de), du substantif **agh** (lait) + **-t** (deuxième élément du circonfixe du genre).), euphorbe ou bien avec **-asif** “rivière” et **aman** “eau”: **asghersif** “peuplier”, Zwawa, **asghar** “bois, arbre” + **asif**; **tamegdaman**, “pl. de pâturage au bord de l’eau”, [**tama**, côté, **eg/ag/ak**, avec, **aman**, eau?]; **tamegdaman** “pl. de pât. au bord de l’eau”, Mtg.; etc.

Verbe +Substantif

tamgersif, plante indéterminée. Laoust la définit comme un “nom de plante poussant dans les rigoles” à Tagountaft. Or, si le nom **asif** est clairement identifié, rien n’est dit sur la première partie du syntagme qui rappelle, pourtant, un autre exemple de nom de plante, plus connu et très souvent cité, composé apparemment de la même racine **mgr** dans le sens de “rencontrer, aller à la rencontre de” et de **aman** “eau”: **taddjaman** [**teğaman**] ? “*Aeluropus littoralis*” (du verbe **eğğ**, laisser, quitter + **aman**, eau?); **talidaman**, “plante aquatique”, A. Ouirra [verbe **ali** “monter” + **aman** “eau”?];

amageraman, aunée visqueuse, *Inula viscosa* (BELL: 93; Trabut 1935: 135; DAL: 490, etc.).

Composés avec tif/tuf/f/af

Les constructions avec base /tff/ dans le sens “être mieux que, surpasser” (Composé ancien puisqu’il est attesté au XI/XII -ème siècle (cf. T&B 2002: 447).) se retrouvent dans plusieurs cas comme: **tifuzzal**; et probablement **teifuzzal** et **tifuzzal**, Zwawa, **tif** + **uzzal**, fer *Taxus baccata* L. “if” (LM: 493; BELL: 503). Le mot **uzzal** apparaît souvent dans le vocabulaire des plantes : **tuzzalt n tzgi**, cytise, Kabylie, **tuzzalt**, “romarin” b. Menacer, **tuzala** id., b. Snous; **tuzzalt**, “frêne”, Aurès, d’une racine **uzzal**, selon Mercier (1905: 85), car le bois de cet arbre est d’une dureté extrême; **tuzelt**, *fraxinus dimorpha*, Sud-Oranais; **tifs(i)ghra**, pour le moment non identifié (“genêt du Sahara” ? *Onobrychis viciaefolia* Scop., T&B: 266); **tif el Hinna** (T&B: 447); **tifelleft**, navet, Zwawa, **if**, meilleur + **lleft**, navet, excellente variété de navet; **tufgho** [**tufaghu**], euphorbe, A. Atta = **tuf** + **agh**, lait (LM: 494); **tufagho** [**tufaghu**], “nom d’une composée à fleur jaune”, A. Atta; **tuftolba** Igr., Ichq., Bugle, ivette, BELL: 277, litt. : “Meilleure que les médecins”. **tufaman**, “pl.

aquatique” (A. Atta, LM: 494) qui devrait donner quelque chose comme “meilleur que l’eau”. Difficile d’en imaginer le sens. Une autre possibilité serait d’y voir la racine des exemples précédents **af**, “trouver / rencontrer” ou bien peut-être, mais moins probable, la corruption d’une préposition **af**, “sur”, dans de très nombreux parlars: “celle qui rencontre/trouve l’eau; celle qui est (flotte ?) sur l’eau”.

Composés avec eilal “herbe”

La Umda da **aṭrīlāl** comme *Ptychotis verticillata* Duby (T&B: 427; BCT- II: 35). Nous avons là un cas intéressant dans la mesure où ce terme est identifié dans la majorité des sources comme d’origine berbère avec la traduction “pied d’oiseau” (R. Basset, BELL, Topper, Aḥmed Issa, Chafik (TOME I, 412) donnent **aḍar waylal** “riḡlu al-ḡurāb” littéralement “pied du cordeau”, etc.). LM (474), en revanche, cherchera ailleurs une étymologie-explication au phytonyme en interprétant le terme **aḍrilal** dans le sens de “herbe” (Ghat) et qu’il retrouve dans **ḍerilel** dans le sens de “pâturage en général” dans le Taitoq. Il en tire la conclusion que l’expression est composée d’un élément final **ill**, pl. **illen**, qui désigne en Ahaggar, “un très beau pâturage, très vert et très abondant”. Le même auteur identifie cette plante comme “astragalus” (LM: 493), tout en l’écrivant d’une autre forme: **aḍreilal**. Ce terme existe par ailleurs dans d’autres langues berbères comme le kabyle par exemple sous la forme **adrilal** mais il est aussi attesté comme **aḍar-waylal** [aḍar waylal] et **aḍariylal** (Boogert 1998: 173 et 185) ainsi que dans de nombreuses sources anciennes. Signalons également, en faveur de cette thèse, la racine /LT/ comme variante possible, utilisée dans le sens “plantes, herbe, parfois légumes verts ou même “bette” (**tilitin**) attestée dans le *Kitāb al-Barbariyya*: **tilitin** “plantes, herbe” (Brugnatelli 2014: 134) comme dans l’exemple: **aman ḡḡ_emm”neṭ tilitin** “Eau où l’on a fait cuire des herbes” (Idem: 131).

Bien entendu, cette issue n’en est pas une, puisque, d’une part, elle n’explique pas la première partie du terme (**aṭar/aḍar**) et, d’autre part, évacue le sens “pied d’oiseau”, relevé par plusieurs auteurs.

- La *Tuḥfat al-Aḥbāb* [TUH] la cite comme **aṭrīlāl** (TUH: 127) **aṭrīlān** *Ptychotis Ammi*. Ce terme, **aṭar ilāl**, dont il reconnaît l’origine berbère et la signification “pied d’oiseau” en s’appuyant sur R. Basset, mais en lui opposant la version de Laoust qui est la seule à ne pas faire référence à la traduction “pied d’oiseau”.

Certains, comme El Ghassānī (Al-Ḥaṭṭābī 1985: 256 et 257), qui y voit une expression utilisée dans la région de Fès, ne citent pas son origine berbère. Les deux variantes **aṭrīlan** et **aṭrīlal** sont identifiées par l’éditeur du manuscrit comme “cerfeuil”, *Garum ammioides* et traduites par **riḡl el ḡhurāb**, litt. “Pied de corbeau”.

C’est peut-être en s’appuyant sur cette version que Chafik (TOME I, 146) reprend l’identification “cerfeuil”, pour **aṭarwaylal**, avec la variante fricative **aḍarwayla**, dont il reconnaît, en revanche, la traduction littérale du berbère “pied d’oiseau”, mais là aussi, sans donner des informations quant à la composition de l’expression.

En revanche, aucun de ces auteurs n’arrive à étayer l’explication étymologique “pied d’oiseau” par la composition du syntagme.

Il est vrai que si la première partie de l'expression **aḍar/aṭar** ne pose aucun problème pour y trouver l'équivalence du sens "pied" en berbère, il en est autrement de la seconde partie de l'expression sur la base /l/ ou /yl/.

De fait dans l'écrasante majorité des parlers berbères, le terme "oiseau" est rendu par une racine de base /gḍḍ/ > **ajḍiḍ**, **agḍiḍ** et le mot "aile" par la base /fr/ et /frw/ dans le sens de "voler", "aile" etc. Pourtant, le touareg nous donne une base sémantique synonyme /yl/, moins connue dans les dialectes du Nord, mais néanmoins attestée dans la variante tamazight de l'Atlas dans le même sens. Taifi (1991: 783) relève une forme verbale **ayll/aylla** "voler, s'envoler" ainsi que le nom **aylalen (wa)** "ailes, gent ailée", mais encore plus évident est Laoust qui relève pour le tachelhit **ailal** "collectif, petits oiseaux" (Laoust 1921: 92). "voler, en parlant d'un oiseau" chez les A. Warain et les Izayan (LM: 471). Dans le même ordre d'idées, Laoust (LM: 493) donne **tameṭṭereilalt** comme "n.p.n. persistante, Ahaggar; composé de **eilal** "herbe" en dialecte Ahaggar. Dans le paragraphe suivant, le même mot est mis en relation avec un sens **tamet**, indéterminé.

De nombreux autres cas demeurent complètement ou en partie sans solution, mais parmi les plus intéressants, signalons quelques éléments pré- ou suffixés.

Composés avec mer-

Peu étudiés et intrigant pour le moment de nombreux chercheurs. Donnée par Nait Zerrad (2002: 355) comme une variante de **ber-**, malheureusement sans plus de détails, ce préfixe apparaît dans une série de noms composés à initiale (en plus de **ber-**) **ba-**, **ma-** ou justement **m(a/e)r-**.

Laoust (1920: 494) donne des exemples comme: **merzizua**, mélisse; **timerzizin** "plante indéterminée", **Merzigiḍen** "grande marguerite", Zemmour; **timermenna**, sureau, **tamerdghaten**, "arbrisseau indéterminé", **timerzuza buḍrar**, *Isatis Djurdjurae*, etc. Ce préfixe provient probablement de la réduction d'un ancien substantif, mais il est jusqu'à présent difficile de faire des hypothèses quant à sa possible signification. Pourtant des indices intéressants confirment sa nature préfixale. Il est à relever par exemple que, à côté de **merzizwa**, cité par Laoust, sont également attestées les formes **tiziwit** et **tizizwit** dans le même sens: "mélisse" (Trabut 1935: 339; Issa 1930: 117).

De même, Traut (1935: 69) donne pour la chicorée, *Cichorium Intybus*, **arhlilou**, **timerzouga**, **mersag**, **timirzagt**... Cet exemple indique que la racine **rzg**, "amer" est le noyau sémantique (verbe: **irzig**, substantif **tarzugi**) autour duquel s'organisent ces termes, qui dans certains cas prennent, semble-t-il, ce préfixe sur base /m-/ et surtout /mer-/ comme dans **amerzagu**, **imirzag** (kab., rifain, etc.).

D'autres noms de plantes présentent le même morphème sans qu'il soit toujours possible de déterminer s'il s'agit du même ou d'un préfixe arabe construit sur la base de la racine /mr/ qui signifie justement également "amer": **marar**, comme dans *mararat aš-ṣaḥāri*, "coloquinte".

Ces termes sont-ils en relation avec d'autres comme kabyle **mernuyet** *Marrubium vulgare* (DAL: 516), plante également amère, ou bien avec des formants

comme **merz-** comme dans l'exemple **Merṣbuqal / lemṣbuqal**, “variété de liseron” (DAL: 518).

Composés avec wa-

À relever également une caractéristique assez courante dans le lexique des plantes: la fréquente préfixation d'un morphème *wa-* aux noms de plantes amazighes. Cette observation avait déjà été faite par Laoust (1920: 508 et *suiv.* qui y voit un figement du démonstratif **wa-** ou **wi-**, attesté dans tous les dialectes amazighs. Elle a été reprise récemment par Brugnatelli (1998), qui élargit néanmoins cet usage à d'autres champs lexicaux: **wabiba / abiba** “moustique”; **wafud/afud** “genoux”; **wagerzam/agerzam** “léopard”; **waru/ari** “joie, bonheur”; **wawter, wawtert** “humérus”/**awtâr** “cuisse” (Rif) etc. Cependant, la frappante sur-représentativité du morphème préfixé **wa-** dans le vocabulaire des plantes demeure toujours inexpiquée.

Déjà, Laoust donnait une liste de 80 noms de plantes qui commencent par ce préfixe (LM: 508–524) et qui comprend autant les noms qui commencent par un formant **wa-X/wi-X** que ceux qui commencent par un **war-X**.

Après Laoust (1920), Marcy (1931) et Brugnatelli (1998) a repris le sujet de la morphologie des noms en **wa-** et l'idée de considérer ce morphème comme un ancien préfixe en les identifiant en cela aux particules démonstratives *wa* et *-a* connues dans la grande majorité des parlers berbères. Il est certain que les arguments de Brugnatelli, qui apporte un nombre de 150 cas de noms dans différents parlers ne manquent pas d'attirer l'attention sur ce phénomène, particulièrement important dans le vocabulaire des plantes.

Or, si ces explications qui mettent en relation le préfixe **wa-** avec un ancien article-démonstratif, déjà signalé par Marcy (1931: 60), Laoust (1920) et repris par Brugnatelli (1998), peuvent paraître plausibles, il n'en est pas ainsi du préfixe **war-**.

En effet, cette catégorie de noms à initiale **war-** n'ont généralement pas de correspondants dénués de ce formant, comme c'est – en revanche – très souvent le cas pour les noms à préfixe **wa-/wi-**. Par ailleurs, ce préfixe rappelle bien sur clairement la particule de la négation berbère **war/ur**: **warnella** “indéterminé, Ill.”; **Waruri**, sureau, Ihn.; **warinsa**, “indéterminé” Mtg.”; **warneger, Osyris alba**, Zwawa (LM: 518).

Ce dernier exemple est traduit par “Celui qui laisse des enfants males, des rejetons”.

Composés avec -enni,-enna, -emmi

Ces formants, dans ce cas, suffixés, demeurent une énigme pour les chercheurs. Ils interviennent dans des termes comme **azukenni**, thym; **iertenni [yertenni]**, “sorte d'arbre à fleurs jaunes”; **amerzgenni/amerzgelli** (indéterminé); **waserkenna**, (feuilles de l'azouka), **égersemmi**, nom d'un arbre; **waserkenna**, feuilles de l'azouka; **wahfenna** (indéterminé); **taferjemma**, “sorte de charbon” (LM: 495–496).

Nous pourrions y ajouter d'autres termes comme: **tibinsert** > **tibis'ennit** (Trabut 1935: 21) *Altaea officinalis* L.; **âzukenni**, **tazukennit** (BELL: 307), *Thymus* divers. . .

Rappelons tout de même que ce suffixe est également attesté dans le système verbal. Nous le retrouvons par exemple en kabyle dans des formes comme **Wweɣnenni**; **yeɣweɣnenni**; **awerennni**, « se tortiller, être tordu, dévié, mal aligné » (DAL: 874); en rifain comme dans **ɣnunni**, “dégringoler” (Cadi, Kaddour [CK]:154) ou plus généralement dans le sens “être rond” (CK: 137) ou dans **karnunni**, “être rond, être bouclé (cheveux, laine)”, (CK: 142, 57); “s’arrondir en boule, sphère” (AR: 401); **tkarnunnuy** (CK: 57, etc.).

Ce suffixe est en relation avec le touareg **mənənnəɖ** et ses dérivés dans le sens « être tordu » (Prasse et alii 2003: 621) et surtout **enned** “enrouler, envelopper, tourner, tordre, être enroulé etc.” que nous retrouvons également dans plusieurs autres dialectes comme le kabyle ou le rifain: **ennaɖ**, “tourner autour, s’enrouler” (Renisio 1932: 419). Le /ɖ/ pouvant être un morphème expressif, comme dans Kab. **šebbed**, “grimper”, tamazight du Maroc central **šubbed**, “grimper” (Nait Zerrad 2002), ce suffixe pourrait ajouter une connotation dans le sens de “ce qui s’arrondit, se tortille, se boucle. . .”, point inimaginable pour une plante. À l’appui de cette hypothèse, relevons que le touareg connaît **mānānaɖ**, que Prasse et alii (2003: 621), donnent comme « esp. de plante (*Cucumis prophetarum* L.) employée comme épice dans d’autres remèdes traditionnels (pour traiter les contusions).

Archéologie linguistique et récupération-reconstruction lexicale

Il est certain que l’étude de ces sources permet de récupérer un nombre appréciable de termes qui ont, soit partiellement, soit complètement disparus, ou bien dont on n’imaginait même pas l’existence en amazigh. La seule analyse des termes contenus dans le manuscrit de la *‘Umda at-tabīb* permet la récupération des termes suivants, perdus dans la plupart des parlers actuels du Nord:

Adigal, “Pastèque” (*Citrullus lanatus*, *Citrullus vulgaris* Schrad.); (L’auteur spécifie cet emploi chez les Almoravides (voir aussi BCT-III: vol. 1, 244).) **agan**, “variété de melon vert”, *Cucumis melo* L. (LM: 419); **ighri** / **taqlilišt**, “asphodèle”, *Asphodelus microcarpus* Viv. pour **aberwaq** (BCT- II: 81); **igg/iğğ/ijj/hejji**, “pistachier térébinthe”, *Pistacia terebentus* (BCT-II: 167, entrée 1119); **tiqidda/tiqda**, “coriandre”, *Coriandrum sativum* (BCT-II: 159); **tabdigha**, “Sang dragon”, *Dracaena draco* (BCT-II: 159); **taffart** ou **tarast** (BCT-I: 149, note 4), probablement corruption de **tafrast** “poireau sauvage” (LM: 493). Trabut (1935: 19) reprend ce terme qu’il redonne, par ailleurs, dans le dialecte arabe sous la forme *firass*. Pour le “poireau cultivé”, le terme **Tarnast** est encore attesté (Trabut 1930: 20; DAL: 594). Un des noms les plus intéressants est sans doute le terme *tabudught*, “coton” [touareg et dans la *teggargrent*, Ouargla: **tabudught**] (BCT-II: 68). Chafik (TOME II, 325) donne **tabɣught** et **tafdught** “vêtement en coton”. Une variante **tabdoq** / **tibdughin** apparaît chez Foucault (1951–1952: I, 30) pour “coton non filé, cotonnier, fruit du cotonnier, coton”. Le parler de Ouargla atteste également une variante **tafduxt** (Delheure 1987: 2), qui renvoie au touareg *tabdoq*, dérivée de la

racine /**fdgh**/; **taghist**, “grenade” *Punica granatum* (BCT-II: 165, 331); **taghesht** (Chafik: TOME I, 229); **taqindawt**, “noix de coco”, *Cocos nucifera* (BCT-II: 166); **tizleft/tazleft**, “la cuscute”, *Cuscuta* sp. (BCT-II: 557); **tifist**, “lin” *Linum usitatissimum* L. Chafik (II, 375) donne **tiffest** (BCT-II: 161), **tifist**; **tifi**, “variétés de pins”, *Pinus pinea*, *P. maritima*, *P. halepensis*; **tafurut** [*tafrut*], iris (*Iris germanica* L., *I. florentina* L. (BCT-II: 165); **tilult**, “câpres”, *Capparis spinosa*; **tigerqesht**, “prunes”, *Prunus armeniaca*; **tiqirut/iqiru**, “carvi”, *Carum carvi*.

Les emprunts

Il existe plusieurs études consacrées aux emprunts berbères à d’autres langues. Le fameux travail de Schuchardt (1918) est certainement la référence par excellence sur le sujet puisque l’auteur aborde les emprunts au latin dans différents domaines dont celui des plantes. Cet ouvrage a été repris par de nombreux travaux sur l’influence du latin sur le berbère.

Laoust reprend à son compte - pour le domaine des plantes - les exemples suivants:

- **tifirest**, “poire commune” < *pirus*; **ulmu**, orme champêtre, *ulmus*; **gernuneš**, “cresson” < *crisonus*; **ifilku**, “fougère aigle” < *Felix, Felicis*; **blitu**, blette, *Atriplex hortensis* (DAL: 26; HL: 115) < *blitum*, mais aussi “chénopode”, *Chenopodium album* L. (BELL: 164); **abellaluz**, “tige d’asphodèle” < *bulbus*; **tinuat** [**tinwat**] (Imettugen, tachelhit), **iten** (Rif), “chêne à glands doux”, “Tan > écorce de chêne vert” du latin < *tannum* qui rappelle l’allemand *Tanne* (LM: 478).

De l’amazigh vers le latin?

Lorsque l’on parle des relations entre l’amazigh et le latin, on ne pense pratiquement jamais que le mouvement pourrait aller dans les deux sens. C’est comme si nous nous trouvions devant un prédéterminisme, une règle non écrite selon laquelle lorsqu’un terme amazigh peut être rapproché du latin, c’est “bien sûr” le latin qui “donne” et l’amazigh qui “reçoit”.

Les chercheurs reprennent souvent les éléments tels quels et attribuent des origines et des étymologies à des langues “prestigieuses” comme le latin ou l’arabe, alors que l’emprunt peut se réaliser dans le sens contraire et passer des provinces amazighes de l’Empire romain vers le latin ou de l’Afrique du Nord vers l’arabe.

Pourtant, les indices nous sont souvent délivrés par les auteurs et les ouvrages latins eux-mêmes, lesquels, dans l’impossibilité, parfois, de trouver une étymologie latine ou grecque à un mot déterminé s’en remettent à des expressions vagues et imprécises comme “latin d’Afrique” ou “origine inconnue” (André 1985).

Ceci est probablement le cas de nombreux termes. Quelques-uns sont déjà connus. *Tabuda*, *tamarix*, *tagantes* mais aussi **taslighwa/siliqua** ont été signalés par H. Schuchardt (1918: 16–27) comme possibles passages du berbère vers le latin.

- **tabuda** > latin *buda*, -ae: “quenouille” ou “la massette” (*Typha* sp.). Le mot est attesté dans ce sens dans la majorité des parlers berbères. Le mot est attesté dans ce

sens dans la majorité des parlers berbères. Le rapport de ce terme avec le latin a été signalé par Colin (1926–1927: 60–61), mais sans référence au sens de l’emprunt. Ce terme est également repris par Bustamante - Tilmatine (1999: 51), qui signalent que ce terme est déjà considéré par André comme d’origine amazigh (note de bas de page 6) et abordent certains sens dans certaines langues berbères de ce terme.).

- **taghundast** > **tagantes**, **-is**: “pyrèthre” *Tanacetum vulgare* L. Le mot dans une variante “**tagantes** (-tis) -is, est donné par André (1985: 255) comme étant d’origine “inconnue”. Il apparaît dans d’autres œuvres de la littérature classique sur les plantes (*’Umda*, *TUH*), ainsi que chez H. Schuchardt (1918: 16), qui voit dans le *ta-* initial la marque du féminin berbère.
- **salghwâ** > **siliqua**, **-ae**: “caroubier”, arabe: **xarrûb šâmî** (*Ceratonia Siliqua* L.). Cette forme, **salghwâ**, citée déjà par la *’Umda* comme berbère est probablement déjà arabisée. La forme amazighe du mot est **taslighwa** avec des variantes comme **tasliugha**, **slighwa**, **tisliwxa** dans le nord et le centre du Maroc (Renisio 1932: 315; Ibañez 1949: 35). Relevons qu’Ibañez, qui avait également repéré l’attestation du nom en berbère et en latin, préconisait un emprunt du berbère au latin (Ibañez 1947: 226–229; Ainsi que 1961: 447–455, et bien d’autres). Etymologiquement, la racine **slgh** existe dans la majorité des dialectes dans le sens de “coller, empâter, agglutiner, matière gluante” ou bien directement “résine” comme c’est le cas de **aselgha** en rifain p. ex. (Renisio 1932: 448).

Relevons d’ailleurs que Schuchardt en parlant du *Quercus aesculus* qui donne en berbère **iškir** (DAL: 86) n’exclut justement pas le fait que le sens de l’emprunt soit contraire, c.-à-d. du berbère vers le latin: “Nun könnte ja *aesculus* aus dem Latein ins Baskische und Berberische, aber ebenso ins Latein aus dem Iberischen oder Libyschen, zu sehr früher Zeit, eingeführt worden sein” (Schuchardt 1918: 18).

La même chose vaut également pour le cas de **akerruŝ** [**akerruŝ**], “chêne vert” et ses variantes comme **axerruŝ**, **ašerruŝ** qui proviendrait du latin *quercus* selon (LM: 506), mais qui selon Schuchardt (1918: 19) pourrait bien avoir fait le chemin contraire: “Das berb. *-kerruŝ* könnte [...] ins Latein übergetreten sein: *cerrus*”.

Par la suite, quelques voix éparses se joindront à ce doute comme c’est le cas de G.S. Colin (1957: 7), qui dès la première page signale que si des travaux, consacrés aux emprunts du berbère aux autres langues existent, il n’en est pas de même du sens inverse, d’où son intérêt pour les mots berbères passés dans le dialecte arabe de Malte. Plus récemment, V. Brugnattelli (1999), revenait sur le sujet dans un article bilan sur les relations du latin avec le berbère.

D’autres termes peuvent donc encore apparaître au fur et à mesure que les études berbères avancent et que l’on envisage les possibilités des deux sens dans les transferts.

En fait tous les termes spécifiés comme “latin d’Afrique” peuvent déjà susciter l’attention du chercheur, comme **gelala**, coloquinte que J. André (1985: 109) donne comme “latin d’Afrique”, ajoutant “cf. arabe *gelala* de même sens”.

Les emprunts d’origine punique (suffixe – im)

Les termes qui se terminent par un suffixe en **–im**, traité par H. Stumme (1899 et 1912), H. Schuchardt (1912), LM (1920: 273), Vycichl (1952) ont été mis en relation

avec l'influence du punique sur le berbère. Selon Schuchardt (1912: 170), ces termes, avec à l'origine une forme du pluriel en **-im** ont été introduites dans le libyque et y ont survécus comme collectifs.

Les mêmes exemples se retrouvent dans les différentes contributions. Laoust (1920: 496) cite quelques termes, comme: **aḡhanim**, roseaux; **aḡarrim** qu'il donne comme non identifié, mais qui pourrait être mis en relation avec le tachelhit Stumme (1899: 228 et 1912: 125) **taḡurrimt** "Grasstoppe" (touffes d'herbe); **aḡessim**, **agerrum** et **tagerrumt** (Zemmour, Maroc), melon vert. L'auteur donne également en kabyle, **Ardrim** et **ardlim** "cerisier sauvage", *Cerasus avium*, (Zouaoua) ainsi qu'avec une variante dentale emphatique; **ar(e)drim**, donné par Schuchardt (1912: 169), comme **arḡelim areḡlim** pour "cerises sauvages" ("wilde Kirschen, Kirschen im allg.") et la forme **tareḡdrimt** pour l'arbre ou le fruit; **aḡazalim**, "oignons", **tehatimt**, "olivier cultivé", ainsi que d'autres termes non-cités par Schuchardt comme **intrim**, "laiteron", **iṭim**, "centaurée", **tarèkimt**, "navet", **agursēlem**, **igurslēmen**, "champignons", (Chaker (1980–1981: 147) inclut ce terme sous les emprunts au latin.) **azarem**, "baie du jujubier", **agultem**, "indéterminé"...

Il est certain que l'influence du punique semble peu discutable. Il est en revanche peut-être moins évident que la seule présence d'un suffixe pourrait justifier l'existence de cette influence punique. En effet, beaucoup d'autres termes non-cités par ces auteurs (Schuchardt, Vycichl, Stumme notamment), présentant les mêmes caractéristiques pourraient, dans ce cas, être ajoutés à la liste. Citons, parmi la liste non exhaustive les cas suivants: **artazim**, *Nitraria tridentata* Desf, type de jasmin (Schuchardt 1912: 169); **ḡardjem**, *Arthrocnemum macrostachyum* (Trabut 1935: 35); **hamahim**, *Ocimum basilicum* (Trabut 1935: 175); **aouchmien** (Trabut 1935: 15); **Acanthyllis numidica**; **azim**, *Launea resedifolia*, *L. quercifolia* (Trabut 1935: 146); **balim**, *Salix*; **bekḡqim**, *Ustilago*; **belsem**, *Tanacetum vulgare* (Trabut 1935: 253); **belsim**, Lens; **bersim**, *Trifolium* (Trabut 1935: 259); **R'ardem [ḡhardem]**, *Zygophyllum album* (Trabut 1935: 273); **terkem**, *Brassica napus*, synonyme **tarekimt**, *lleft* en arabe dialectal, navet, (Trabut 1935: 48). etc...

Conclusion

Le vocabulaire ancien des plantes tel que reflété dans les œuvres anciennes peuvent servir, sans aucun doute, pour donner de la profondeur historique à grand nombre de phénomènes linguistiques autour des langues amazighes. L'étude des noms des plantes dans les sources anciennes comme la 'Umda confirment les résultats obtenus de l'étude de textes plus fournis des manuscrits anciens (comme Brugnatelli 2014: 127–142) et ce à plusieurs niveaux d'analyse comme la variation phonétique (e.g. spirantisation des occlusives, variation phonétique, apophonie...), la syntaxe avec l'existence de l'opposition d'état. Du point de vue lexical, le vocabulaire phytomique se distingue assez clairement du reste du lexique par un recours très prononcé au phénomène de la composition comme moyen de création terminologique et, bien entendu, par

l'existence d'un grand nombre d'archaïsmes, qui pourraient s'avérer très utiles aujourd'hui. Un travail plus profond des sources anciennes et du vocabulaire des plantes en général pourrait donc générer la collecte d'un grand nombre de noms de plantes aujourd'hui disparus ou inusités.

La persistance des noms de plantes et leur caractère parfois endémique, rend naturel et compréhensible le fait que des peuples étrangers les désignent par leur nom local. Il en a été ainsi également en Afrique du Nord où des noms locaux ont pu passer à l'autre côté de la Méditerranée même si, souvent, on oublie que les "petites" langues peuvent aussi donner aux "grandes". Certains termes sont facilement reconnaissables, par exemple, à leur formant *ta-* comme marque du féminin. Schuchardt (1918: 16) a été un des premiers à donner une liste des noms de plantes passés au latin. La langue des plantes recèle également des indices sur les emprunts d'origine punique en berbère.

Toutes ces données, une fois collectées, systématisées et étudiées peuvent apporter une aide non négligeable aux efforts actuels en vue de standardiser et de développer la langue amazighe en vue de se lancer dans ce gigantesque chantier que constitue, cette phase de passage à l'écrit avec tout ce que cela implique comme travail de création lexicale et d'équipement en instruments linguistiques. Ces quelques observations permettront, peut-être, de contribuer à faire l'état des lieux et de mesurer les faiblesses en termes de lexicologie et lexicographie dans un champ sémantique spécifique: celui de la terminologie des plantes et, à partir de là, tenter de dégager certaines mesures d'urgence.

Face à un déficit néologique qui apparaît difficilement récupérable, deux mesures s'imposent. La réalisation dans un premier temps d'un inventaire tant au niveau intra- que supradialectal du lexique amazigh des plantes. Cet inventaire apparaît comme une nécessité et une priorité en raison des fonctions de complémentarité qui reviennent aux différentes langues amazighes. Ensuite, une recherche approfondie et diversifiée qui engloberait à la fois les sources écrites arabes pertinentes de l'époque médiévale et un minutieux travail de récupération des termes amazighs dispersés dans des travaux spécialisés d'origines diverses, devraient permettre – dans une seconde phase – de compléter et affermir les résultats de la première et ainsi, pouvoir aboutir à un produit qui puisse répondre dans une proportion acceptable aux besoins fondamentaux de l'heure, ou du moins faire face - en tous cas dans un premier temps – à la situation actuelle. Eu égard aux problèmes énoncés ci-dessus, il semble évident que l'inventaire ne peut se faire sur la base de racine "porteuse de sens commun". La variation phonétique, lexicale, l'imprécision du champ spécifique de dénomination auront vite raison des efforts dans ce sens. En fait, le seul dénominateur commun entre les différentes dénominations demeure le signifié, la référence, donc la plante. Or, une désignation exacte du signifié est fournie par son nom d'identification latine: la dénomination binomiale, qui semble pour le moment bien être la seule entrée lexicale raisonnable autour de laquelle pourraient s'articuler les différentes variantes dialectales et les commentaires correspondants. Une fois réalisé, ce travail pourrait ensuite servir de référence et de base pour l'élaboration de lexique dans les différentes langues amazighes ou - pourquoi pas – dans quelconque des variantes régionales respectives.

Bibliographie consultée

- Al-Ḥaṭṭābī al-ʿarabī M (ed) (1985) *Ḥadīqat al-azhār fī māhiyyat al-ʿuṣab wa al-ʿuqqār*. Dār el Gharb al islāmī, Beyrouth
- Al-Ḥaṭṭābī al-ʿarabī M (1990) Abū l-Ḥayr al-Iṣbīlī: *ʿUmdat al-Tabīb fī maʿrifat an-nabāt*. qaddama la-hu wa-ḥaqqāqa-hu wa-ʿaāda tartībā-hu Muḥammad al-ʿArabī al-Ḥaṭṭābī, 2 vols. Maṭbūʿāt ākādīmiyyat al-Mamlaka al-Maghribiyya, Rabat
- André J (1985) Les noms de plantes dans la Rome antique. Les belles Lettres, Paris
- Asin Palacios M (1943) Glosario de voces romances registradas por un botánico anónimo hispano-musulmán (siglos XI–XII). CSIC-Escuelas de Estudios Árabes de Madrid y Granada, Madrid-Granada
- Basset R (1899) Les noms berbères des plantes dans le traité des simples d'Ibn el Beitar. *Giornale della Società Asiatica Italiana* XII: 53–66
- Bellakhdar J (1997) La Pharmacopée marocaine traditionnelle. Ibiss Press [BELL], Paris
- Bellakhdar J (2010) Contribution à l'étude de la *ʿumdat aṭ-ṭabīb* d'abulḥayr al-iṣbīlī: commentaires à propos de quelques items et propositions d'élucidation. *Al-Andalus Magreb* 17: 9–38
- Brugnatelli V (1998) La morphologie des mots berbères en w-. Considérations diachroniques. In: Elmedlaoui M, Gafaiti S, Saa F (eds) Actes du 1^o Congrès Chamito-sémitique de Fès, 12–13 mars 1997, pp 51–67
- Brugnatelli V (1999) I prestiti latini in berbero: un bilancio. In: Lamberti M, Tonelli L (eds) Papers from 9th Italian Meeting of Afro-Asiatic (Hamito-Semitic) Linguistics Trieste, April 23–24, 1998, pp 325–332
- Brugnatelli V (2014) Typology of Eastern Medieval Berber. *Language Typology and Universals, Berber in Typological Perspective*, STUF 67(1):127–142
- Bustamante J and Tilmatine M (1999) El léxico amazige contenido en la *ʿUmdat aṭ-ṭabīb*. *El Andalus-Magreb* 7: 43–64
- Bustamante J, Corriente F and Tilmatine M (eds) (2004/2007/2010) *Kitābu ʿUmdati t-tabīb fī maʿrifati nnabāt likulli labīb*. CSIC-CSIC Press, Madrid
- Cadi K (1987) *Système verbal du rifain. Forme et sens*. SELAF, Paris
- Carabaza Bravo JM (1995) *El Kitāb al-Nabāt* de Abū l-Jayr. In: Vázquez de Benito C, Manzano Rodríguez MA (eds) *Actas XVI Congreso UEAI. AECl-CSIC-UEAI*, Salamanca, pp 101–115
- Chafik M (1993/1996/2000) *Al-muʿḡam al-ʿarabī-al-amazīghī*. Maṭbūʿāt ākādīmiyyat al-Mamlaka al-Maghribiyya, Rabat
- Chaker S (1980–1981) La situation linguistique dans le Maghreb antique: le berbère face aux idiomes extérieurs. *Libyca* 28/29: 135–152
- Colin GS (1926–1927) *Etymologies maghribines*. *Hesperis* 6/7: 55–82, 85–102
- Colin GS (1957) Mots 'berbères' dans le dialecte arabe de Malte. In: *Mémorial André Basset* (1895–1956). A. Maisonneuve, Paris, pp 7–16
- Dallet J-M (1963) *Petite botanique populaire*. Centre d'études berbères, Algérie
- Dallet J-M (1982) *Dictionnaire kabyle-français: Parler des At Mangellat*. SELAF, Algérie/Paris
- de Foucauld Ch (1951–1952) *Dictionnaire touareg – français*, vols I–IV. Imprimerie Nationale, Paris
- Delheure J (1987) *Agerraw n iwalen teggargrent-taṛumit/Dictionnaire ouargli-français*. SELAF/CNRS, Paris
- Dozy R (1881) *Supplément aux dictionnaires arabes*, vol. Premier. E.J. Brill, Leiden
- Hanoteau A, Letourneux A (1892–1893) *La Kabylie et les coutumes kabyles*, 2nd edn. L'Imprimerie Nationale, Paris
- Ibañez E (1947) Voces hispano-latinas en el dialecto rifeño. 5: 365–381
- Ibañez E (1949) *Diccionario español- rifeño*. Instituto de Estudios Africanos, Madrid
- Ibañez E (1961) Voces latinas en el dialecto bereber del Rif. *Orbis* 10: 447–455
- Issa A (1930) *Dictionnaire des noms de plantes en latin, français, anglais et arabe*. Gouvernement Press, Cairo
- Kossmann M (1997) *Grammaire du parler de Figuig. Maroc oriental*. Peeters, Paris

- Laoust E (1920) Mots et Choses berbères. Notes de linguistique et d'ethnographie (dialecte du Maroc). Challamel, Paris. [Reprint 1983, Société Marocaine d'Édition, Rabat]
- Laoust E (1921) Cours de berbère marocain. Dialecte du Sous du Haut et de l'Anti Atlas. Challamel, Paris
- Marcy G (1931) Essai d'une théorie générale de la morphologie berbère. *Hespéris* XII(fasc. I):50–90
- Mercier G (1905) Le nom de plantes en dialecte chaouia de l'Aurès. In: XIVème Congrès International des Orientalistes, Alger, t. 2, 4ème section. Ernest Leroux, Paris, pp 79–92
- Moïnfar MD (1988) Patte-de-chat et œil-de-rossignol. Le corps et les membres dans le vocabulaire botanique persan. Société d'Ethnologie, Nanterre
- Naït-Zerrad K (2002) Les morphèmes expressifs en berbère. In: Naït-Zerrad K (ed) Articles de linguistique berbère. Mémorial Werner Vycichl. L'Harmattan, Paris, pp 349–372
- Prasse K-G, Ghoubeïd A and Ghabdouane M (2003) Dictionnaire touareg-français (Niger). Museum Tusculanum Press, University of Copenhagen, Copenhagen
- Renaud, H. P. J., and Colin, G. S. (Éds.) (1934). *Tuhfat al-aḥbāb*. Glossaire de matière médicale marocaine. Paris: Geuthner [Institut des Hautes Etudes Marocaines, tome XXIV].
- Renisio A (1932) Etudes sur les dialectes berbères des Beni Iznassen, du Rif et des Sehnhadja de Sraïr. Grammaire, textes et lexique. Leroux, Paris
- Salmon G (1906) “Sur quelques noms de plantes en arabe et en berbère”, suivi de: “Les Achchabins de Fes”. In: Archives Marocaines, Mission scientifique du Maroc, tome VIII. Ernest Leroux, Paris, pp 1–98
- Schuchardt H (1912) Zu den berberischen Substantiven auf –im. *Zeitschrift für die Kunde des Morgenlandes*, BD XXVI:163–170
- Schuchardt H (1918) Die romanischen Lehnwörter im Berberischen. Alfred Hölder, Wien
- Stumme H (1899) Handbuch des Schilḥischen von Tazerwalt. Grammatik-Lesestücke-Gespräche. Glossar. Leipzig: J.C. Hinrichs'sche Buchhandlung.
- Stumme H (1912) Gedanken über Libysch-phönizische Anklänge. *Zeitschrift für Assyriologie*, BD 27:121–128
- Taifi M (1991) Dictionnaire tamazight-français (Parlers du Maroc central). L'Harmattan, Paris
- Tilmatine M and Bustamante J (2002) La fitonimia amazige en la ‘umdat aṭ-ṭabīb. *Al-Andalus-Magreb* 8/9: 413–462
- Topper U (1999) Wortsammlung der Tier- u. Pflanzennamen im Taschelheyt-Berberischen. *Almogaren* XXX: 229–249
- Trabut L (1935) Flore du Nord de l'Afrique. Répertoire des noms indigènes des plantes spontanées, cultivées et utilisées dans le nord de l'Afrique. N.P., Alger
- van den Boogert N (1989) Some notes on Maghribi script. *Manuscripts of the Middle East* 4: 30–43
- van den Boogert N (1998) “La révélation des énigmes”. *Lexiques arabo-berbères des XVII^e et XVIII^e siècles*. Travaux et Documents de l'Iremam, 19
- Vycichl W (1952) Punischer Spracheinfluss im Berberischen. *J Near East Stud* XI(3): 198–204

Index

A

- Ablaut, 37, 292–294, 300–304
 Absence de l'opposition verbo-nominale, 417
 Achaïsme phonétique, 287, 288
 Acoustics
 analyses, 59
 correlates, 5, 115
 feature of gemination, 4, 5
 silence, 7
 timing patterns, 7, 22
 Actif, 411, 422
 Action noun, 31, 39, 46, 168, 169, 637–656
 Adjectives, 238, 242, 302, 335, 352, 367, 368,
 370, 458, 469, 473, 477, 483, 485–487,
 489, 499, 500, 505–507, 511, 523–544,
 638, 641–643, 653, 654, 656
Ad 'non-realised particle' (Figuig), 203,
 317–329, 464, 484, 547–567, 621, 627,
 662, 669, 706
 Adverb, 428, 431
 Adverbial, 432, 434–440
 Afroasiatic, 26, 28, 89, 155, 167, 519, 572,
 616, 618
 Afro-Asiatic, 6, 61, 154, 231, 291–305, 386,
 392, 395, 468, 617
 Afro-asiatique, 408, 416, 418
 Agentive noun, 28, 168, 368
 Agglutination, 418
 Akerḱur « tas de pierres, endroit sec et pierreux
 », 688
 Akkadian, 28, 294, 296, 300–305
 Al-Andalus et Afrique du Nord, 142, 712, 715
Ald 'until' (Figuig), 550
 Algeria, 8, 162, 257–262, 337, 339, 342, 343,
 386–388, 390, 393, 395–400, 459–463,
 465, 472, 519, 521, 620, 621, 627
 Amazigh
 alphabet, 393, 396–398
 cultural movement, 399
 culture, 197, 258, 394, 397
 language, 61, 153–163, 221, 224, 234,
 388, 391, 393, 395, 397–399, 444, 453,
 497, 616
 population, 399
 Amazighe langue, 140
 Ancien, 275, 280–282, 284, 417, 687, 688, 690,
 691, 706, 707, 709, 710, 714
 Aorist, 29, 31, 36, 39, 42, 168, 292, 293, 295,
 303, 305, 308–310, 312–316, 318,
 320–332, 362–366, 368, 369, 371, 379,
 380, 531, 534, 537, 547–567, 639,
 640, 654
 Aoriste sans préverbe, 283
 Apodosis, 426, 427, 432–440, 550, 553
 Apodosis aorist, 550
 Apophony (apophonic classes), 308, 318, 323,
 364–368, 374, 542
 Apprenant, 130
 Apprentissage, 130, 136, 143–147
 Apprise langue, 140
 Approche, 136, 139, 144, 146, 405, 406, 412,
 414, 415, 689, 695
 Arab conquest of Morocco, 390
 Arabic, 5, 8, 56, 57, 60–63, 76, 77, 79, 80, 98,
 110, 114, 154–162, 166, 167, 169–171,
 175, 177–179, 183–187, 196, 200,
 203–205, 213–219, 221, 222, 224, 225,
 231–237, 239, 244, 245, 249, 252, 257,
 259, 262–267, 294, 297, 298, 302–304,
 313, 326, 331, 340, 346, 354, 385–400,
 426, 463, 477, 478, 496, 499, 517–522,
 525–531, 533–539, 541–544, 560, 564,
 566, 601, 616, 622, 633, 638, 644, 648,
 651, 652, 663, 664, 666, 669, 676
 loans, 155, 157, 264, 313, 346, 390–393,
 496, 536
 Archaïsme, 274–277, 279, 282, 284, 288
 Archaïsme de 1^{er} degré, 276

- Archaïsme de 2^e degré, 279
 Archaïsme de 3^e degré, 276, 277
 Archaïsme de 4^e degré, 276
 Archaïsme fossile, 276, 287, 288
 Archaïsme lexical/lexico-sémantique, 275
 Archaïsme linguistique, 273
 Archaïsme morphologique, 287, 417
 Archaïsme néologique, 276
 Archaïsme syntaxique, 282, 287
 Articulation, 5, 26–28, 35, 37, 60, 97, 98, 110–112, 174
 Articulatory contact extents, 7
 Aspect, 6, 30, 34, 35, 39, 58, 61, 62, 79, 87, 108, 125, 170, 203, 206, 224, 241, 243, 252, 259, 260, 265, 275, 276, 283, 292, 293, 295–298, 313, 342, 374, 403, 426, 431, 433–440, 457, 465, 494, 511, 531, 548–552, 559, 616, 618, 633, 639, 649, 662, 671, 703, 705, 706
 Aspectual system, 292–293, 301, 302, 304, 305
 Assimilated geminates, 114
 Assimilation *ad + aoriste*, 276, 283
 Assimilation réciproque ou amalgame, 278
 Assimilation régressive par assourdissement, 278
 Atomic distribution, 589
 Atomicity, 577, 586–587, 589
 Auditory priming, 170, 171
 Autonome langue, 124, 127–129, 131, 132
 Autonomie linguistique, 128, 131
 Autosegmental phonology, 106, 108, 601
 Autosegmental tiers, 671
 Awjila (Awj.), 156, 319, 322, 332, 619–622
 AXB discrimination test, 68
 Ayer Tuareg, 319, 324, 460, 622
 Ayt Atta, 311, 317, 330, 556
 Ayt Atta Central Moroccan Berber, 551
 Ayt Sadden Central Moroccan Berber, 562
 Ayt Seghrushen (Central Morocco), 552, 564
 Azayar « plaine sèche », 695
- B**
- Base bi-consonantique, 685, 688
 Base radicale, 684
 Beja (Old Cushitic), 292, 303, 304
 Bengali, 4, 5, 110, 111, 114
 Berber
 causatif, 406, 407, 421
 varieties, 156, 159–161, 203, 235, 309–312, 314–317, 319, 320, 322, 324, 430, 433–435, 438, 469, 522, 527, 530, 544, 549, 619, 621, 622, 624
- Berbère
 actuel, 410, 411, 415–418, 420, 422
 consonantal root, 171, 667, 671
 moderne, 287, 406, 408, 410, 411, 414, 416, 417, 419–421
 pré-moderne, 419
 Berbère/Latin/Punique, 61, 156, 161, 170, 173, 214, 232, 233, 235–237, 241, 243–245, 293, 388, 393, 395, 396, 518–521, 663, 667, 671, 679, 712–715
 Bilingualism, 217, 218, 392, 396
 Bilinguisme, 128–130
 Blocages des études syntaxiques égyptien, 415, 421
 Burst amplitude, 108
- C**
- Caractère officiel, 127, 141
 Catégorie morphologique, 406, 684
 Categories, 6, 10, 18–20, 22, 57, 62, 95, 112, 115, 168, 169, 173, 184, 238, 239, 242, 243, 263, 293, 300, 303, 304, 308, 374, 375, 409, 416, 427, 429, 430, 434, 438, 453, 486, 494, 501, 506, 507, 510, 513, 550, 576, 628, 667, 684, 700
 Causative, 39, 174, 318, 352, 362–367, 374, 375, 380, 430, 572, 639, 690
 Central Moroccan Berber (CMB), 310, 315–317, 319, 321–323, 325, 327–329, 332, 551, 555, 557, 562
 Central vowels, 62, 63, 307–310, 316–318, 324, 326, 329
 Chained aorist, 547, 550–552, 555–557, 559–567
 Champ sémantique, 684, 685, 715
 Champs morphosémantiques, 684, 685
 Changement sémantique, 687
 Circulaire, 135, 141
 Classical Arabic, 63, 213–216, 303, 304, 478
 Clauses, 96, 295, 298, 305, 342, 375, 381, 426–428, 432–436, 439, 440, 457, 464, 466, 468, 469, 474, 475, 477, 480, 505, 538, 539, 543, 548, 550, 552–555, 562, 567, 618, 627, 633
 Closed class, 444, 446, 451
 Closure duration, 4, 5, 7, 9, 10, 108, 109, 112, 113
 Codification, 133, 135, 136, 393
 Collective readings, 574, 575
 Colloquial Arabic, 213, 215, 222, 387–393, 396, 400, 664
 Comitative preposition, 347, 625

- Common Semito-Berber, 302, 303
 Comparaison inter-dialectale, 686, 687, 689, 690
 Comparative/typological approach to language reconstruction, 304
 Comparison, 10, 27, 32, 50, 74–77, 108, 113, 161–162, 175, 182, 214, 222, 230, 232, 236–238, 243, 336, 338, 340, 342, 349, 353, 371, 379, 380, 397, 426–429, 434, 437, 439, 519, 562, 584, 616, 651, 667
 Composition, 418, 452, 495, 699, 701, 706–708, 714
 Concatenated geminates, 114
 Conception bassetienne, 404, 405, 411
 Conception pré-bassetienne, 404–407
 Concessive, 426, 439, 440
 Concomitant correlates, 112
 Conditional, 425–440, 551, 553
 Conditionality, 427, 434
 Conjugation, 293–294, 302, 304, 308, 319, 430, 525, 528–534, 543, 544, 638
 Conjunction, 27, 34, 298, 427, 431–440, 617, 625, 631
 Conscience linguistique, 131, 132
 Consonantal closure, 7, 9–12, 16, 18, 19
 Consonantal occlusion, 7
 Consonant frequency, 448
 Consonant permeability, 45–48, 50
 Consonant-vowel (CV)
 phonology, 106
 syllable, 10, 15, 17–22, 182
 Consonant-vowel interactions, 25–51
 Conspiracy, 27, 32, 48–51
 Constituents, 41, 94, 144, 410, 418, 428, 434, 470, 477, 601–603, 679
 Constraint conjunction, 27, 34
 Construction, 139, 170, 240–243, 245, 247–251, 274, 279, 280, 288, 336, 340, 342, 345, 347–349, 352, 353, 372, 411, 412, 417, 420, 421, 426–428, 430, 432, 439–440, 455, 464, 465, 468, 469, 477, 485, 487, 494, 495, 498, 500, 502, 510, 511, 527, 535, 539–542, 549, 550, 552–555, 557–563, 623, 627–630, 633, 638, 687, 705, 707
 Construction accusative, 420, 421
 Contrastive focus, 630, 633
 Contrastive topic, 630, 633
 Co-occurrence restrictions, 29, 30, 32–34
 Copules d'existence *d* et *eg*, 280, 288
 Corpus linguistics, 444
 Corpus parémiologique, 273–289
 Corse langue, 127–131, 133, 136–139
 Couchitique, 417, 418
 Counterfactual, 429, 437, 438
 Counterfactuality, 427, 440
 Crosslinguistic, 427, 458, 459, 487
 Cultural identity, 259, 597
 Curricula, 141, 144, 218, 252
 Cypriot Greek, 110
- D**
Daj mi 'every time that' (Figuig), 553–554
 Decomposition theory, 178, 183, 185
 Défini, 125, 132, 281
 Déformation, 406, 407, 409–411, 415, 420–422, 702
 Degrés d'archaïsation, 274
 Deixonne, 124, 127, 129
 Dérivation, 91, 93–95, 178, 183, 318, 322, 362, 364, 366, 404, 421, 460, 495–497, 499, 501, 503–505, 507–510, 512, 604–606, 608, 610, 651, 663, 664, 668, 671, 678, 695
 Désuet, 275
 Diachronic, 274, 351–354, 362, 372, 430, 438, 461, 476, 487, 528, 597, 639
 Diachronically, 362, 438, 476, 599
 Diachronie, 274, 275, 686, 696
 Dialecte, 127–132, 134, 135, 138, 139, 145, 257, 259, 705, 706, 709–711, 713
 Diathèse, 405, 410, 411, 418
 Diathetic system, 298–300
 Diffusion, 127, 133, 137, 139
 Diphthong, 519, 520
 Direct speech, 558, 563, 566
 Disparu, 275, 276, 282, 706, 711, 715
 Dissimilation, 26–30, 32, 34, 35, 37, 38, 43–45, 48–51, 174, 325, 326
 Distinctive features, 106, 107
 Distributive morpheme, 576
 Distributive numerals, 572–575, 587, 589, 590, 594
 Distributive readings, 572–578, 580–589, 591–594
 Distributivity, 572–576, 584, 588, 589, 594
 Dorsal, 26–30, 34, 449
 Double articulation, 28
 Double passif, 407, 411
 Duration, 3–22, 66, 105–116, 170, 174, 549, 558, 576
 Duration ratio, 115, 116

E

Ecole, 126, 130, 397
 Écriture système d'éducation bilingue, 135
 Education, 62, 63, 125, 135, 141, 197, 198, 204, 205, 213–219, 222–224, 230, 231, 233–235, 241, 258–260, 264, 266, 267, 387–389, 391, 393, 394, 396–398, 400
 Egypt, 217, 257, 386, 390, 459, 518, 616
 Egyptian, 154, 198, 294, 302, 304, 616
 Élaboration linguistique, 129, 131, 135
 Electropalatography, 109
 Emergence of the Unmarked, 31, 36, 43, 49–50
 Emphasis, 26, 89, 97–99, 178, 224, 234, 237, 250, 252, 257, 618, 619, 623, 633
 Emphatics, 65, 69, 75, 307, 448
 Enfant, 126, 139, 143, 337, 703, 710
 Enseignée langue, 129, 136, 139, 140
 Enseignement, 127, 130, 136, 139, 140, 142, 144, 145, 147
 Epenthesis, 30–32, 39, 40, 48, 61, 600, 670, 674
 Époque médiévale, 715
 Error rate, 176, 181
 Estonian, 114
 État d'annexion, 281–283, 409
 État libre, 281, 282
 Event-distributive readings, 572–575, 578, 580–582, 584–589, 591–593
 Event plurality, 573, 576, 577, 588–590, 593
 Event plurality marker, 573, 588
 Events, 9, 198, 200, 218, 244, 292, 297–298, 370, 372, 427–429, 435–438, 548, 549, 551, 553–555, 557–560, 562–567, 571–594, 618, 628, 629
 Evolution phonique, 686
 Exhaustive distribution, 577, 587, 589
 Exhaustivity, 577, 586–587, 589
 Experience in story telling, 562

F

Facilitatory effect test, 171, 177–179, 182
 Factive, 404
 Faithfulness constraints, 35, 37, 41
 Fake geminates, 114, 115
 Famille, 123–149, 284, 416, 684, 686–688, 694
 Famille des mots, 684
 Fast speaking rate, 6, 8, 18, 19
 Feature classes, 41
 Figuig (fig.), 203, 317–329, 332, 464, 484, 547–567, 621, 627, 662, 669, 706

Finnish, 111, 115, 477–479
 Flanking vowels, 7, 11, 13
 Focalization, 619, 627
 Focus particles, 631–632
 Formation lexicale, 686, 692, 693
 Forme dérivée du verbe, 284, 404, 405
 Forme simple du verbe, 284, 404, 421
 Framed Aorist, 550–552, 555, 567
 Free Aorist, 550
 Frequency, 56, 58, 171, 173, 182, 243, 246, 443–453, 526, 549, 551, 552, 557, 562, 648, 655
 Frequency norms, 443–453

G

Geminate ambiguity, 106
 Geminates, 3–22, 27, 33, 89, 103–117, 168, 248, 300, 307, 310, 314, 324, 362, 448–450, 453, 531, 602, 609, 639, 649–651, 654, 668, 669, 672, 674–676, 678
 Geminate voiced stops, 7, 10, 12–18, 113
 Geminate voiceless stops, 105, 111
 Gemination, 4–6, 8, 19, 20, 22, 33, 39, 89, 104–106, 108, 110–117, 185, 248, 294, 323, 325, 362, 367, 368, 374, 380, 381, 609–610, 639, 649–652, 668–670, 674–676, 679, 680
 Gender agreement, 487, 618, 621
 Generative phonology, 48, 105
 Ghadames (Ghd.), 278, 308–315, 319, 321–325, 327–329, 420, 551, 620, 621, 655
 Ghomara, 155, 159–161, 203, 323, 417, 525, 526, 530, 541, 544, 551, 620–622, 624, 632, 638
 Glottophagie, 131
 Grammaticalized lexical auxiliaries, 297
 Graphie, 137–138, 140, 142, 143

H

Hamitic-Semitic, 418
 Hassaniya Arabic, 215, 521
 Heavy Verb (or Heavy verbs), 307–332
 Hégémonie processus d' généralisation idiome, 129, 131, 132, 134
 Hilalian invasion, 521
 Hindi, 110, 482
 Hypothetical, 427–429, 432, 437, 438, 440, 486
 Hypotheticality, 426, 427, 434–440

I

- Identity avoidance, 27, 30, 34–38, 43, 44, 47, 48, 50
- Ilelli/tilelli « libre/liberté », 284, 287
- Ili* ‘be’ (Figuig), 295, 297, 431, 432, 534, 549, 555
- Imperative, 168, 250, 313–315, 428, 503, 550, 551, 556, 559–561, 567, 639, 649, 654
- Imperfect, 305, 321, 430, 431, 433, 435–438
- Imperfective, 89, 104, 107, 108, 168, 292–305, 308, 314–318, 320–332, 374, 531, 548–551, 553, 555–558, 563, 567, 572, 639, 656
- Implicature, 436
- Indéfini, 281
- independent personal pronouns, 615–633
- Indicateurs linguistiques, 128, 132
- Individuation, 128, 131, 132
- Information structure, 627–629
- Inhibitory effect, 171, 179, 182, 186
- Ini* ‘say’ (Figuig), 314, 555, 563, 566
- Instrument noun, 168
- Intégration, 127, 197, 205, 230, 252, 308, 392, 393, 399, 414, 521
- Intensive, 31, 39, 40, 42, 241, 252, 362–365, 367, 369, 371, 380, 430–432, 435
- Intercompréhension, 144
- Internal variation, 433, 619
- Intersegmental data, 11, 13
- Intersegmental parameters, 12–14
- Intersegmental results, 11–15
- Invariance, 115–116
- IRCAM, 61, 140, 167, 197, 219, 220, 230, 234, 258, 394, 395, 397, 398
- Italian, 110, 115, 160, 354, 477, 478, 483, 520, 663–665
- Iterativeness, 432
- Iwelmeden Tuareg, 319
- I* ‘when’ (Figuig), 434

J

- Japanese, 3–22, 59, 69, 115, 161, 330, 354
- Juxtaposition, 248, 250, 433, 624, 695, 706

K

- Kabyle, 155, 157, 159, 273–289, 294, 295, 298–302, 304, 308, 311, 312, 315–317, 319–323, 325–332, 339–340, 347, 351, 352, 390, 393, 399, 406, 408, 417, 418, 420, 551, 556, 562, 572, 597–611, 616, 620, 621, 627, 630, 649, 685–694, 705, 706, 708, 709, 711, 714

Kabylie

- extrême orientale, 276, 277, 287
- occidentale, 276, 277, 280
- orientale, 280, 282

L

- Labial, 27, 28, 30, 116, 168, 174, 327, 364, 449
- Labialised consonants, 25–51
- Language
- attitudes, 211–226, 252, 259–262
 - change, 386, 390–393, 479
 - contact, 153, 231, 262–263, 265, 386, 390–392
 - processing, 166, 169–172, 444
 - revitalization, 399
 - teaching, 57, 78–80, 234
 - varieties, 213, 220
- Langue berbère, 277, 694, 706
- Langue nationale, 124, 125, 140, 142
- La voyelle initiale, 275, 281, 285
- Length, 5, 10, 105, 111, 117, 154, 156, 244, 247, 249, 251, 328, 373, 378, 520, 601–605
- Lexical decision task, 171, 172, 179, 182, 185, 186, 452
- Lexical geminates, 114–116
- Lexical representation, 35, 171, 172, 187, 378, 445, 453, 676
- Lexicon, 62, 63, 104, 114, 155, 162, 166, 167, 170–177, 183–187, 215, 390, 391, 444, 446, 450–453, 496, 503, 506, 507, 676
- Lexicon organization, 187
- Lexique géographique, 688
- Linear-mixed-effects regression model, 175
- Linguistic diversity, 203
- Locality, 45, 46, 48
- Locuteurs du 1^{er} âge, 276
- Locuteurs du 2^e âge, 276
- Locuteurs du 3^e âge, 276

M

- Maghreb, 79, 206, 264, 386, 387, 390, 391, 399, 400, 518, 521, 701, 702
- Malayalam, 5, 110
- Mali Tuareg, 319, 322, 323, 327
- Manner class, 449
- Manuel, 133, 137, 143–147
- Manuels scolaires, 144–147
- Markedness constraints, 34, 35, 37, 41, 44, 49, 50
- Mauri, Simone, 311, 548, 550–552, 555–557
- Measurements of absolute durations, 19–20

- Mediopassive, 298
 Mental lexicon, 170–176, 187
 Middleweight verb (or middleweight verbs),
 316–318, 321
 Minoration, 123–149
 Minorée langue, 124
 Minorisation, 124, 125, 136
 Minoritaire langue, 124, 126
mi ‘when’ (Figuig), 553–554
 Monolingualism, 58, 216, 387, 393
 Moroccan Arabic, 8, 56, 57, 61, 62, 77, 79, 80,
 98, 110, 170, 171, 179, 184, 185, 221,
 231, 233, 245, 389, 392, 517–522, 525,
 666, 669
 Morocco, 8, 26, 61–63, 66, 167, 195–198,
 200–204, 206, 207, 212–225, 230–234,
 236, 257–262, 346, 386–388, 390,
 393–400, 430, 434, 440, 444, 459,
 461–465, 484, 493, 518, 520–544,
 547–567, 620–622, 632
 Morphological geminates, 104
 Morphologically, 40, 104, 166, 173, 176–178,
 180, 184–188, 292, 353, 368, 372, 392,
 430, 438, 446, 476, 495, 496, 499, 539,
 541, 543, 549, 577
 Morphology, 28–30, 34, 35, 39, 40, 56, 61, 108,
 162, 166, 168, 169, 171–173, 177–181,
 184–186, 307, 350, 353, 354, 371, 375,
 450–452, 459–461, 496, 504, 506, 520,
 526, 548, 593, 619, 621, 623, 639,
 655, 671
 Morpho-syntactic class, 451
 Multilingualism, 56, 233, 400
- N**
 Natural class, 449
 Negation, 276, 280, 374, 375, 380, 419, 438,
 525, 536–538, 543, 548, 567, 710
 Negative, 59, 168, 175, 212, 213, 220, 225,
 252, 258, 259, 262, 265, 293, 295, 362,
 374, 398, 438, 473, 548–550, 567, 639,
 640, 652, 655, 656
 Nyenbalu « vendetta/assassinat par balle ? »,
 286, 287
 Neutralization, 6, 16, 519
 Neutralized, 6, 519
 Non-orientation ou disponibilité du prédicat,
 408–412
 Non-verbal predications, 622, 624
 Normalisation orthographique, 133, 136
 Normalization, 69
 Normal speaking rate, 6, 8, 11–19
- Norme graphique, 136–139, 142
 Norme plurielle, 133, 135
 Noyau sémique, 685, 688, 690, 695
n-tuples, 445
 Numeral reduplication, 571–594
- O**
 Officialisation, 123–149
 Officielle langue, 124, 127, 140, 141, 145
 Omotique, 417, 418
 Optimality theory (OT), 27, 31, 57
 Ouargla (Ouargli), 316, 317, 320, 321,
 323–332, 553, 621, 688, 711
- P**
 Paléologisme, 275, 276, 284, 287
 Paradigme personnel suffixal, 417
 Parenté linguistique, 128, 129
 Parlers, 130, 140, 277, 278, 280, 338, 339, 408,
 414, 417, 418, 420, 686, 687, 693, 706,
 711
 Participant-distributive reading, 573, 574, 576,
 578, 580–582, 584, 589
 Particles, 295–298, 305, 337, 342, 346–348,
 350, 353, 369, 371, 372, 392, 431–433,
 435–438, 457, 528, 537, 543, 548–550,
 553, 572, 622, 624, 631, 633, 663
 Passif, 404, 405, 407, 411, 421, 422
 Passive, 30, 31, 33, 34, 45, 47, 130, 280, 293,
 299, 300, 302, 318, 404, 405, 408, 410,
 411, 416, 431, 526, 639, 671, 676
 Patrimoine national, 140
 Perception, 6, 57–59, 65, 66, 68, 69, 73, 75–80,
 105, 106, 112–114, 116, 260
 Perfect, 79, 155, 292–304, 313, 323, 331, 351,
 430, 431, 435–440, 519, 561
 Perfective, 91, 104, 107, 108, 168, 292, 295,
 300–305, 308–312, 315, 316, 318,
 320–332, 362–374, 380–382, 530, 531,
 534, 536, 537, 539, 548–551, 553–556,
 558, 559, 562–567, 572, 639, 640, 652,
 656, 671
 Pharyngeals, 97, 168, 245, 449, 669
 Phonetics, 6, 57–60, 62, 64, 76, 88–90,
 105–108, 111, 112, 115, 116, 168,
 183, 249, 251, 274, 339, 348, 350,
 392, 393, 395, 445, 453, 497, 567,
 572, 602–605
 Phonological
 contrast, 5–8, 113
 geminates, 104

- Phonologie
 diachronique, 692
- Phonology, 28, 30, 32, 33, 48–51, 57–63,
 87–99, 103–117, 155, 162, 178–181,
 186, 187, 189, 214, 236, 369, 392, 519,
 521, 601–603
- Phrasal verbs, 305
- Phytnymie amazighe, 702
- Plain vowels, 308–310, 316, 318
- Pluractionality, 572, 573, 576–577, 594
- Pluractional markers, 572, 573, 576, 581, 588,
 590
- Pluractionals, 572, 573, 576–577, 581, 588,
 590, 593, 594
- Plural events, 573, 592
- Plurality requirement, 577, 587–588
- Plurinorme, 129, 134, 135, 138, 139
- Politique linguistique, 124, 126, 130
- Polynomie, 129, 131, 134, 135
- Polynomique langue, 129–131, 133–136, 138, 145
- Possessive pronouns, 630
- Possibility, 45, 48, 51, 112, 337, 354, 362, 392,
 396, 429, 432, 435, 436, 512, 526, 527,
 560, 617
- Postverbal position, 627, 629
- Pragmatically, 440
- Pragmatic functions, 616, 617, 623, 628, 663,
 664, 669, 677–679
- Pragmatics, 426–428, 433, 440, 476, 548, 549,
 552, 616, 617, 623, 628, 629, 632, 664,
 669, 677, 679
- Prédicat d'existence, 411, 417, 419
- Presupposition(s), 439
- Preterite, 29, 31, 34, 37, 38
- Preverbal position, 628
- Primaire, 130, 140, 141, 143–147, 335
- Primary correlate, 108–113
- Primary labial consonants, 28, 30
- Priming condition, 172, 173, 175, 176, 180, 181
- Priming effect, 170–173, 176–179, 181–186
- Priming experiment, 166, 171, 179, 186
- Probability, 11, 426, 427, 436, 437, 439, 440,
 447, 453
- Production didactique, 135, 136, 142–144
- Profondeur historique et continuité du lexique
 des plantes, 701, 702, 710
- Programmes pédagogiques, 144, 146
- Pronouns, 240, 251, 336, 345, 346, 348, 352,
 353, 482, 502, 519, 615–633
- Proposition, 137, 342, 407, 427, 428, 433, 439,
 440, 501, 619, 703
- Prosodic contour, 630
- Protasis, 426, 427, 432–438, 440
- Proto-Berber, 155, 156, 291–305, 307–332,
 416–420, 687
- Proto-berbère, 415–420, 687
- Protosémantisme, 694
- Prototype, 438
- Prototypical, 439, 458, 482, 483, 487
- Proverbe, 275–279, 281–283, 285–288
- Psycholinguistic, 167, 187, 444, 452, 618
- Q**
Qeccuc meccuc, 284, 285, 287
Q'im 'stay' (Figuig), 549
- R**
 Random intercepts model, 175, 176
 Reaction time, 113, 166, 175, 176, 178, 181,
 183, 185
 Reconstruction linguistique, 684, 686, 689
 Rection, 418
 Redoublement consonantique, 692
 Réduplication/redoublement, 294, 295, 298,
 301, 317, 332, 418, 571–594, 665, 670,
 672, 675, 678, 679
 Reference, 28, 60, 95, 128, 137, 138, 141, 144,
 145, 230, 231, 234, 236, 240, 250, 296,
 297, 318, 336, 345, 354, 370, 371, 404,
 405, 426, 427, 429, 432, 436, 437, 440,
 480, 531, 534–539, 543, 549, 551, 554,
 617, 618, 623, 624, 628, 632, 666, 706,
 708, 712, 713, 715
 Referents, 133, 336, 436, 539, 616, 619,
 624–626, 628, 630, 677
 Régionale langue, 124, 125, 128
 Relative clause, 295, 342, 538, 539,
 543, 627
 Relative stability, 20–22
 Relative timing, 20, 22
 Relative values, 10, 15–20
 Résultatif, 412–414
 Root, 35–37, 39, 47, 166–169, 171–173,
 176–188, 239, 241, 267, 312, 316, 317,
 319, 321–325, 327, 328, 332, 349, 352,
 353, 363–365, 367, 371, 374, 377–380,
 382, 418, 446, 452, 505, 526, 530,
 598–601, 604, 605, 607–611, 638, 639,
 649–651, 655, 663, 666, 667, 669–676,
 678, 679
 Root effect, 177
 Round, 26–28, 30, 32–37, 41, 43–45, 48, 50,
 311, 312, 328
 Rural Arabic, 62, 389, 392

S

Sad 'future particle' (Figuig), 431
 Sauvegarde, 124–127, 140, 148
 Sauvetage, 126
 Secondary articulation, 26–28, 35, 37, 97, 98
 Secondary correlates, 110–112
 Secondary labial consonants, 34
 Secondary labialization, 34, 449
 Segment frequency, 452
 Semantics
 features, 95, 172, 173, 177, 185, 187
 relatedness, 171, 172, 177, 178, 184, 187
 Sémantique et structure morphologique du
 vocabulaire des plantes/Sources
 anciennes et modernes, 701–702
 Sémitique, 404–406, 412–416, 418, 420–422,
 684
 Senhaja (Snh.), 319, 328, 417, 523–544, 619,
 621
 Singleton voiced stops, 13, 14, 17, 18
 Singleton voiceless stops, 12, 14–16
 Siwi, 155, 157–160, 330, 459, 598, 616, 617,
 623, 624, 627, 628, 631–633
 Société, 130, 147, 684, 687, 691, 692, 694
 Sociolinguistique, 125–129, 131, 132, 134,
 137, 138
 Spanish Arabic, 519, 520, 663
 Spatial-distributive readings, 573, 576, 577,
 580, 585
 Spatiotemporal separation of events, 576
 Speech priming, 165–189
 Speech rate, 6, 9, 10, 18, 20–22, 105, 115
 Speech signal, 6, 7
 Spirantization, 115, 348
 Standard deviation, 69, 176, 181
 Standardization, 219, 230, 231, 233–235, 252,
 386, 393, 400
 Stative, 293, 294, 296, 297, 300–304, 308, 412,
 413, 526, 528, 530, 544, 548, 551, 555,
 638–640
 Structuration du lexique, 684, 685
 Structures formelles, 684, 685
 Structures sémantiques, 684
 Stylème, 276
 Subevents, 579, 580, 582, 584, 585, 587, 591,
 592, 594
 Subjunctive, 431
 Subordinate, 298, 427, 432, 435, 436, 550, 552,
 553, 555, 562, 567
 Subordinator, 439, 550, 553, 554
 Supports didactiques, 144, 145
 Supraliminal priming/non-masked
 priming, 170
 Surface variant, 446

Survivance, 275, 276, 287
 Synchronie, 275
 Syntaxe du verbe, 405–407, 415, 422
 Syntaxe ergative, 419, 420, 422
 Système normé, 139

T

Tamaghit Tuareg, 319
 Tamazight, 26, 61, 93, 140, 154, 155, 159, 167,
 196, 218, 233, 258, 278, 279, 287, 292,
 293, 295–299, 301, 353, 386, 397, 418,
 430, 436, 459, 464, 465, 479, 525, 621,
 709, 711
 Taqbaylit Berber, 6–8, 11, 13, 15–18, 22,
 362–364, 373
 Taqbaylit language, 20, 155, 461–467, 473,
 477, 479–481, 484, 571–594
 Taqbaylit numerals, 462, 573–576
 Tarifit Berber, 4, 6, 8, 11, 13, 15–18, 205
 Tarifiyt, 155, 159–161, 163, 263, 551
 Tashelhiyt (tashl.), 295–297, 301, 308, 317,
 319, 332, 551, 556
 Tashlhit, 25–51, 55–80, 165–189, 656, 671
 Tashlhiyt, 87, 89, 90, 92, 97, 98, 103–117, 155,
 312, 315, 316, 320–332, 433, 443–453,
 493–513
 Tchadique, 416–418
 Temporal-distributive readings, 576, 581, 594
 Tense, 5, 78, 105–107, 111, 112, 263, 292, 296,
 305, 370, 428, 431, 440, 469, 667
 Tier conflation, 114
 Token/type distinction, 451, 452
 Topicalization, 551, 552, 627–629
 Tradition berbérissante, 403, 406, 407, 409–412,
 415–417, 421, 422
 Traité de botanique médiévale, 700
 Transcription du berbère, 288
 Tuareg, 155–157, 161, 162, 233, 294, 298, 300,
 302, 308–332, 336, 337, 340, 343–346,
 348, 351–354, 386, 459, 460, 519, 522,
 553, 556, 598, 619, 621, 622, 632, 650,
 651, 655, 705
 Tuareg (Adagh), 550
 Tuareg (Ahaggar), 161, 551, 552
 Tuareg (Tu.), 310, 313, 319, 321, 332
 Tudalt Tuareg, 319, 323
 Typology, 426–429, 458, 459, 466, 476–477,
 616–618

U

Unification, 133, 136
 Universal quantification, 573, 586, 587, 589

- Urban Arabic/Universal quantifier, 572–579, 586–589
- Usure phonétique tigemmi « maison », 686
- Uvulars, 8, 26, 27, 30, 104, 168, 449, 522
- V**
- V1, 5, 9, 11–14, 405, 412, 416–421
- V2, 9, 404, 412, 416, 419–421, 552
- Valeur passive du verbe simple, 405, 408, 411, 416
- Variability, 5, 22, 105, 607
- Variation linguistique, 129, 133–135
- Variétés linguistiques, 127, 129, 686
- Velar, 8, 26–28, 30, 33, 35, 111, 168, 174, 298, 324, 325, 351, 449, 552, 650
- Verbalization, 405, 413, 414
- Verbes de procès, 406, 410, 415, 416, 419, 420
- Verbes d'état, 404–406, 408, 410, 412–420
- Verbes statifs/Vocabulaire botanique, 405, 406, 412, 416, 417, 419, 420
- Verb system, 263, 291–305, 315, 365
- Vieilli, 275
- Vieillissant, 275
- Vocalism, 310–318, 321, 332, 344, 519, 521, 522
- Voice onset time (VOT), 5, 7, 9, 14–18, 20, 59, 60, 111, 113
- Voicing class, 449
- Vowel
- copying, 26, 27, 32, 37–50
 - durations, 5, 9, 112, 115
 - frequency, 444, 446–451, 453
 - length, 156, 520
 - pattern, 309, 638–640, 648, 655, 656
 - slot, 309–311
 - systems, 64, 75, 76, 80, 519
- VTT, 7, 9, 10, 14–16, 20
- W**
- West Semitic, 302–305
- Word frequency, 444, 446, 447, 452, 453
- Z**
- Zenaga (Mauritania), 347, 459, 637–656
- Zenaga (Zng.), 155–157, 308–311, 313, 314, 319–327, 332, 347–349, 352, 459, 528, 530, 531, 551, 553, 567, 598, 620, 621, 637–656
- Zipfian distribution, 446